



UNIVERSITA' DELLA CALABRIA

Dipartimento di Studi Umanistici

Scuola Dottorale Internazionale di Studi Umanistici

Indirizzo

Filosofia della comunicazione
e dello spettacolo: teoria e storia dei linguaggi

Settore Scientifico Disciplinare M – STO/08

CICLO

XXVI

TITOLO TESI

Strategie *Linked Data* per gli Archivi Istituzionali *Open Access*.

Per una migliore comunicazione dei contenuti di ricerca.

Direttore: *Ch.mo Prof. Roberto De Gaetano*

Firma

Coordinatore dell'indirizzo.

Relatore:

Ch.mo Prof. Daniele Gambarara

Firma

Correlatore:

Prof. ssa Anna Rovella

Firma

Dottoranda: *Dott./ssa Iryna Solodovnik*

Firma

ABSTRACT

Lo scopo di questo lavoro è mettere in luce:

1. La rilevanza del movimento internazionale *Open Access* che si prefigge l'obiettivo di sostenere e promuovere le strategie per la disseminazione del materiale di ricerca con l'appoggio di tecnologie di rete e politiche *Open Access*.

2. La prima strategia dell'*Open Access* nota come "*Green Road*" ("via verde", gli Archivi aperti *Open Access*), focalizzandosi su uno dei suoi canali comunicativi per il deposito (l'auto-archiviazione) e la diffusione del materiale di ricerca – gli Archivi Istituzionali (*Institutional Repositories, IR*) basati sulle piattaforme tecnologiche *Open Archives*. I medesimi Archivi saranno presentati entro la cornice di:

- a) alcune linee di indirizzo, soffermandoci sulla presentazione di una serie di Raccomandazioni, Linee guida e i servizi per la *governance* dei contenuti di ricerca "*SurPlus Open Archive*" e "*U-GOV Ricerca*";
- b) valutazione della ricerca a livello istituzionale e nazionale;
- c) questioni di etica collegate al rispetto di proprietà intellettuale e contratti editoriali;
- d) alcune considerazioni relative alla conservazione dei contenuti digitali e all'allestimento di un *Data and Process Management Plan* per programmare e monitorare l'intero *ciclo di vita* dei contenuti digitali;
- e) requisiti di *Trusted Digital Repositories* che si riferiscono agli Archivi digitali sostenibili e affidabili, i presupposti essenziali per la certificazione della loro qualità.

3. Un'indagine su un campione di 23 Archivi digitali (nazionali e internazionali) a sostegno della comunicazione e del Libero Accesso ai contenuti di ricerca in rete.

4. Alcune buone pratiche a supporto del "*profilo tecnico*" dell'Archivio dei contenuti digitali di ricerca che riguardano gli aspetti di:

- a) Piattaforme Software;
- b) Metadati;
- c) Controllo di autorità di valori di metadati;
- d) Associazione di metadati con valori semantici;
- e) Conservazione a lungo termine;
- f) Digitalizzazione;
- g) Nuovi paradigmi di ricerca e navigazione.

L'allestimento dei medesimi aspetti è assemblato in un *ToolKit* che potrebbe essere utile per programmare e monitorare diverse funzionalità di un Archivio Istituzionale.

Il valore aggiunto della Tesi consiste nello sviluppo di uno schema di metadati per la descrizione (catalogazione) dei materiali di ricerca nell'Archivio Istituzionale, prendendo in considerazione:

- 1) "Linee guida per la creazione e la gestione di metadati nei *Repository* Istituzionali" (CRUI, Italia, 2012);
- 2) "LODE-BD (*Linked Open Data-enabled Bibliographical Data*) *Recommendations*" (AIMS, 2012). Le Raccomandazioni "LODE-BD" si prefiggono l'obiettivo di abilitare i metadati – riferiti agli indici descrittivi dei contenuti digitali di ricerca - di acquistare le caratteristiche pertinenti alla tecnologia per la pubblicazione dei contenuti sul Web Semantico - *Linked Data*. A tal scopo si propone di qualificare le proprietà di metadati proposte dalle Linee CRUI con i valori estensibili di vocabolari controllati, schemi semantici (thesauri, classificazioni) e schemi di metadati, al fine di abilitare i collegamenti tra metadati e unità informative aperte (pubblicate con formati e Licenze aperte) correlate nel contenuto e contesto sul Web Semantico. L'ambizione più grande dello spazio rappresentativo fondamentale del Web Semantico - *Linked Open Data (LOD) Cloud* - è quella di costruire una vera e propria enciclopedia di concetti e di relazioni che sia condivisa tra e usata da diversi soggetti produttori dei contenuti digitali.

L'affiancamento di metadati negli Archivi Istituzionali dalle strategie di codifica "LODE-BD" può avere esiti alquanto positivi per:

- a) la qualificazione del profilo semantico delle proprietà dei medesimi metadati;
- b) il potenziamento della comunicazione dei contenuti depositati basata sulla creazione di accessi multifocali ai contenuti correlati (entro e fuori l'Archivio) abilitati dalle logiche di relazioni di *Linked Data*;
- c) il favorimento della condivisione di unità informative tra diversi soggetti produttori dei contenuti digitali, avvantaggiando così lo sviluppo di dinamiche collaborative del materiale di ricerca a livello nazionale e transfrontaliero in rete e, quindi, rafforzando l'efficacia e l'efficienza della ricerca stessa.

ABSTRACT

The purpose of this paper is to highlight:

1. The importance of the international Open Access movement aiming to support and promote strategies for dissemination of research material relying on internet technologies and Open Access policies.
2. A strategy of the Open Access movement known as "Green Road" referring to Open Access Repositories (Archives) comprising Disciplinary and Institutional Open Access Repositories. Particular attention will be paid to the Institutional Open Access Repositories (IRs). These are digital platforms conform to Open Archives Initiative (OAI) requirements and are designed for deposit (self-archiving) and for dissemination of research material produced within certain Institutions such as Universities and Research Centers.

The attention on IRs will be focused within the framework of:

- a) some Guidelines, Recommendations as well as digital services such as "SurPlus Open Archive" and "U-GOV-Ricerca" for governance of research contents;
 - b) evaluation of research at the institutional and national levels;
 - c) ethical issues concerning respect of intellectual property and publishing contracts;
 - d) considerations relating to preservation of digital information resources;
 - e) Data Management Plan, which is important to program and monitor the entire life cycle of digital content objects (digital resources and their metadata, research data);
 - f) requirements for Trusted Digital Repositories (TDR) referring to sustainable and reliable digital Archives having essential requirements for certification of their quality focused on long-term preservation.
3. A survey of a sample regarding 23 digital Archives (national and international) supporting communication of digital contents and Free (Open) Access to research material by means of web technologies.
 4. Some good practices regarding the support of 'technical profile' of digital Archives, which cover the following seven aspects:
 - a) Software platforms;
 - b) Metadata;
 - c) Authority control of metadata values;
 - d) Associating metadata with semantic values;
 - e) Storage and preservation;
 - f) Digitizing;
 - g) Searching and browsing paradigms.

These and other aspects can be assembled in a ToolKit that might be useful to program, manage and update various IR's functionalities.

The added value of the present work is to develop a metadata schema describing IR's research materials taking particularly into account:

- 1) "Guidelines for the creation and management of metadata in Institutional Repositories" (CRUI, Italy, 2012) and
- 2) "LODE-BD (Linked Open Data-enabled Bibliographical Data) Recommendations" (AIMS, 2012). This document proposes encoding strategies enabling metadata values (describing digital information resources) to acquire characteristics of Linked Data constituent entities, through assigning to certain metadata properties "non-literal" values expressed (identified) by web addresses <[http://URIs](#)>. These last should be obtain from authority documents such as controlled vocabularies, standards, authoritative semantic schemes whose meaning can be interpreted within Semantic Web space.

The biggest ambition of the high promising Semantic Web technology- which is Linked Data - is to build a virtual global encyclopedia based on a big amount of related concepts published on the Web according to certain rules by various agents operating with their content on the web.

By modeling IR's metadata according to "LODE-BD" encoding strategies different positive outcomes can be expected, such as:

- a) enriching of semantic profiles of metadata;
- b) improving cross-communication of digital contents based on creating multifocal accesses to correlated digital contents (within and outside Repositories) enabled by relational logic of Linked Data technology;
- c) better sharing of information entities exposed by different agents on the web favoring the development of collaborative dynamics of research material at a global level in internet, enriching research context and increasing efficiency of cross-domain research.

STRUTTURA DELLA TESI

La prospettiva tematica della Tesi si articola in tre parti.

PARTE PRIMA. Il **Capitolo Primo** introduce il lettore alla conoscenza dei presupposti, delle importanti dichiarazioni, delle politiche, delle iniziative e delle attività a sostegno del movimento internazionale *Open Access*, cui obiettivo principale è quello di favorire una migliore disseminazione dei materiali di ricerca.

Il **Capitolo Secondo**, a partire dall'introduzione del progetto *Open Archive Initiative* a supporto delle piattaforme tecnologiche "Archivi aperti", mette in luce una delle due strategie *Open Access* "*Green Road*" ("via verde") che riguarda la disseminazione dei contenuti digitali di ricerca tramite gli Archivi aperti (Disciplinari e Istituzionali) *Open Access*. Una particolare attenzione è riservata agli Archivi Istituzionali, intorno a cui vengono indagate alcune questioni relative a:

- a) una serie di linee di indirizzo e servizi per la loro implementazione e gestione;
- b) la valutazione dei prodotti di ricerca;
- c) l'etica nel rispetto di contratti editoriali e proprietà intellettuale;
- d) la conservazione a lungo termine dei contenuti digitali.

Ci si focalizza sull'importanza della stesura delle Politiche (*Policy*) a sostegno degli Archivi Istituzionali, il presupposto importante per lo sviluppo di *Data Management and Process Plan* per ottimizzare la programmazione dell'intero ciclo di vita dei contenuti ospitati nei medesimi Archivi. Il Capitolo si conclude con la trattazione di una realtà pratica - *Trusted Digital Repositories*, che si riferiscono al profilo di Archivi digitali che offrono i servizi sicuri, affidabili e sostenibili nel tempo e rappresentano una premessa fondamentale per certificare la qualità di Archivi digitali.

PARTE SECONDA. Il **Capitolo Terzo** espone un'indagine su un campione di 23 infrastrutture nazionali e internazionali a sostegno della comunicazione del sapere in rete e del Libero Accesso ai contenuti di ricerca. Il **Capitolo Quarto** presenta un *ToolKit* basato sullo stato dell'arte di alcuni modelli, progetti e buone pratiche a supporto del "profilo tecnico" degli Archivi digitali di ricerca riferendosi ai seguenti aspetti: "Piattaforme Software" – "Metadati" – "Controllo di autorità di valori di metadati" – "Associazione di metadati con valori semantici" – "Conservazione a lungo termine" – "Digitalizzazione" – "Nuovi paradigmi di ricerca e navigazione".

PARTE TERZA. Il **Capitolo Quinto** è dedicato alla presentazione del valore aggiunto della Tesi, in particolare all'introduzione di uno schema di metadati a supporto descrittivo dei contenuti di ricerca negli Archivi Istituzionali *Open Access*. Il medesimo schema sarà sviluppato in conformità con "Linee Guida per la creazione e la gestione di metadati nei *Repository Istituzionali*" (CRUI, Italia, 2012) e "*LODE-BD (Linked Open Data-Enabled Bibliographic Data) Recommendations*" (AIMS, 2012), ed è volto a potenziare la comunicazione dei contenuti digitali entro e fuori le piattaforme dei contenuti di ricerca, con la proiezione sull'ambiente dei contenuti digitali nel Web Semantico.

Alla fine la Tesi è allestita con le **APPENDICI** in cui vengono descritti alcuni approcci a supporto della rappresentazione, della gestione e dell'interoperabilità dei contenuti digitali.

In particolare, nell'**APPENDICE 1** vengono presentate le strategie di metadati (dati bibliografici, indici descrittivi, marcatori, dati sui dati) nella ricognizione di aspetti strutturali e semantici di contenuti digitali. Successivamente, vengono introdotti alcuni linguaggi e soluzioni tecnologiche del Web Semantico: *Resource Description Framework (RDF)*, *Ontology Web Language (OWL)*, *Linked Data (LD)* o *Web 3.0*, *Topic Maps (TM)* cui obiettivo è quello di formalizzare le relazioni semantiche tra diverse classi di unità informative (es. metadati, documenti, collezioni, vocabolari e loro parti) in formato digitale.

È messo in luce il lato pratico del linguaggio grafico di *logica* programmabile *Unified Modeling Language (UML)* che si rivela molto utile per strutturare ontologie di unità informative a supporto di funzionalità e processi comunicativi di un sistema (es. documento, collezione, Software). Viene, inoltre, brevemente descritto il valore di tecnologie e pratiche del *Social Web (Web 2.0)* per una cooperazione costruttiva dei contenuti digitali.

L'**APPENDICE 2** offre uno spazio per presentare alcuni approcci a supporto della programmazione e dell'interoperabilità dei contenuti digitali di ricerca con cui potrebbe essere affiancato un Archivio Istituzionale. Il lettore è introdotto alla conoscenza del progetto dell'Archivio digitale di ricerca *Open Access VOA3R* che aggrega i contenuti da diversi Archivi Istituzionali e offre le funzionalità avanzate di ricerca e navigazione dei contenuti di ricerca, basandosi sulle potenzialità del Thesaurus AGROVOC disponibile online in modalità *Linked Open Data* e cui valori sono identificati tramite URIs (*Uniform Resource Identifiers*) basati sul protocollo web http.

STRUCTURE OF THE THESIS

The thematic perspective of the present work is divided into three parts.

PART I. The **First Chapter** introduces the reader to the knowledge of current situation, statements, policies, initiatives and activities in support of the international Open Access movement, whose main objective is to encourage better dissemination of research materials.

The **Second Chapter** begins with the introduction of Open Archive Initiative project in support of digital Open Archives platforms conform to certain Open Standards. Afterwards the Chapter highlights one of the two Open Access strategies - "Green Road" concerning dissemination of digital research content through (Disciplinary and Institutional) Open Access Repositories implemented on Open Archives software platforms. Particular attention will be paid to the Institutional Repositories (Archives) or IRs, around which some issues will be described. These last are related to:

- a) a set of Guidelines, Recommendations and services for the implementation and management of IRs and their contents;
- b) evaluation of research products;
- c) ethics in respect of publishing contracts and intellectual property;
- d) long-term preservation of digital contents;
- e) importance of drafting Policies in support of IRs' functionalities. These Policies are necessary prerequisites for the development of Data Management and Process Plan through which it would be easier to optimize and to plan the entire life cycle of digital contents deposited in IRs.
- f) Trusted Digital Repositories (TDR) referring to the profile of digital Archives offering services enable to provide reliable, sustainable and durable access to digital contents.

PART II. The **Third Chapter** presents a survey of a sample of 23 national and international digital infrastructures (mostly Institutional Repositories), supporting and promoting communication and dissemination of knowledge with research value world-wide through the internet.

The **Fourth Chapter** presents a ToolKit based on information reporting state of the art of some models, projects and best practices in support of technical profile of digital Archives. This ToolKit assembles the following aspects:

- a) Software platforms;
- b) Metadata;
- c) Authority control of metadata values;
- d) Associating metadata with semantic values;
- e) Storage and preservation;
- f) Digitizing;
- g) Paradigms of searching and browsing.

PART III. The **Fifth Chapter** is devoted to the presentation of the added value of the thesis. Here, some approaches for the creation of qualitative metadata supporting the description of research digital content objects in Institutional Repositories are described. Afterwards the metadata schema is presented, developed in accordance with "Guidelines for the creation and management of metadata in Institutional Repositories" (CRUI, Italy, 2012) and "LODE-BD (Linked Open Data-Enabled Bibliographic Data) Recommendations" (AIMS, 2012). This schema is intended to enhance communication (interaction) of digital information contents within and outside Institutional Repositories platforms, focusing on interoperability of digital contents in the Semantic Web environment.

At the end of the thesis there are **APPENDIXES (A, B)** describing some approaches supporting organization, management, representation and interoperability of digital contents. Particularly, **APPENDIX A** introduces metadata (conceived as data about data, descriptive indexes,

bibliographic data) strategies focused on organization of structural, functional and semantic aspects of digital contents. Afterwards, some languages and technological solutions of the Semantic Web will be introduced, such as Resource Description Framework (RDF), Ontology Web Language (OWL), Linked Data (LD) or Web of Data (Web 3.0), Topic Maps (TM), whose objective is to formulate semantic relationships among different entities or classes of information units (e.g. information resources and their parts, metadata on their support, collection of information resources, vocabularies, semantic schemes and their constitutive units) which can be understood, shared and enriched in digital environment.

Subsequently, the practical side of a graphic language - Unified Modeling Language (UML) is highlighted. UML is considered to be a useful tool to structure and represent Ontologies of information units. UML supports programmable logic of different functionalities and processes in regard to information structures, systems and platforms (e.g. document, collection, Software). Then the value of technologies and practices of Social Web (Web 2.0) is described. This Web is an important tool to strengthen communication of information contents provided by different communities of users on the web.

APPENDIX B offers a space to present some approaches supporting conceptual organization, management and interoperability of digital content objects in digital research environments. Here also an Open Access Repository for aggregated digital research contents - VOA3R - is introduced. This Repository provides the advanced features for digital content search and navigation, taking advantage of potential from implemented AGROVOC Thesaurus, available online as Linked Open Data and whose literal values are supported by “non-literal” ones, identified by URIs (Uniform Recourse Identifiers) based on http internet protocol.

RINGRAZIAMENTI

L'esperienza del Dottorato, ricca di stimoli e motivazioni, mi ha dato modo di crescere, agendo contemporaneamente sull'acquisizione delle nuove competenze personali e professionali, per cui non posso che trarre un bilancio positivo.

Vorrei ringraziare il Gruppo di ricerca operativa a sostegno dell'Archivio Istituzionale dell'Università di Vienna (Austria) PHAIDRA (*Permanent Hosting, Archiving and Indexing of Digital Resources and Assets*), la cui attività ho potuto costantemente seguire durante il mio periodo di ricerca all'estero (ottobre 2012 – ottobre 2013).

In particolare, ringrazio il gruppo di lavoro PHAIDRALab: un luogo di incontro di professionisti dell'informazione e della comunicazione (ICT), che contribuiscono allo sviluppo delle nuove funzionalità e della performatività del digitale nell'Archivio PHAIDRA.

Il “*know-how*” e la dimensione collaborativa del PHAIDRALab vengono trasferiti tramite una ampia rete di cooperazione tra diversi Archivi e Progetti a livello internazionale (sito PHAIDRA *international*).

Partecipare alla vita di diverse comunità di ricerca è un'esperienza unica e indimenticabile, attraverso cui si riceve il bagaglio di conoscenze unico, in un interscambio profondo fra sé e i gruppi di appartenenza, il risultato di cui è una migliore organizzazione delle proprie conoscenze, della formazione, della comunicazione e la contribuzione alla ricerca stessa.

ACRONIMI

- AIP** *Archival Information Package.*
Pacchetto di Informazioni per l'Archiviazione, definito dall'OAIS (*Open Archival Information System Standard*) e destinato alla conservazione a lungo termine:
<<http://www.iasa-web.org/tc04/archival-information-package-aip>>
- API** *Application Programming Interface.*
Ogni insieme di procedure disponibili al programmatore, di solito raggruppate a formare un set di strumenti specifici per l'espletamento di un determinato compito all'interno di un certo programma. Spesso con tale termine si intendono le *Librerie Software* disponibili in un certo linguaggio di programmazione:
<http://it.wikipedia.org/wiki/Application_programming_interface>
- BIBO** *Bibliographic Ontology* (Ontologia bibliografica).
Un'Ontologia utilizzata per descrivere le entità bibliografiche come libri e riviste sul Web Semantico:
<<http://bibliontology.com/>>
- CERIF** *Common European Research Information Format.*
Un formato che assicura che le informazioni presenti in fonti di ricerca differenti (sia comunitarie che nazionali) siano accessibili in un formato comune e paragonabile:
<<http://cordis.europa.eu/cerif/src/about.htm>>
- CMS** *Content Management System (CMS).*
Uno strumento usato dai *webmaster* per gestire un sito web evitando di dover programmare dinamicamente l'intero Software lato server che gestisce il medesimo sito: <http://it.wikipedia.org/wiki/Content_Management_System>
- CRIS** *Current Research Information Systems.*
Sistemi informativi per la ricerca sono i sistemi software sia per tracciare e coordinare progetti, attività, gruppi, competenze e risultati delle proprie attività scientifiche, sia per rispondere all'esigenza di razionalizzare la gestione della ricerca, a fronte di un progressivo calo degli investimenti. CRIS adottati dalle diverse università europee sono diventati il principale collettore delle attività e dei dati riguardanti la Ricerca:
<http://en.wikipedia.org/wiki/Current_Research_Information_System>
- DCAM** *DCMI Abstract Model. Un modello astratto usato dai metadati Dublin Core:*
<<http://dublincore.org/documents/abstract-model/>>
- DCMI** *Dublin Core Metadata Initiative.* Un'organizzazione non-profit coinvolta nello sviluppo degli Standard di metadati interoperabili: <<http://dublincore.org/>>
- DCAP** *Dublin Core Application Profile.* Nell'ambito della DCMI, un profilo applicativo è la dichiarazione dei "metadata terms" che un'organizzazione, una risorsa informativa, un'applicazione o una comunità di utenti usa nel suo schema di metadati: <<http://dublincore.org/documents/usageguide/glossary.shtml#dcaps>>
- DIP** *Dissemination Information Package.*
Pacchetto di Informazioni per la Distribuzione trasferito dall'OAIS (*Open Archival Information System Standard*) all'utente in base ad una richiesta di accesso:
<<http://www.iasa-web.org/tc04/formats-and-dissemination-information-packages-dip>>
- Drupal** È un *Content Management Framework, Content Management System (CMS)* modulare scritto in linguaggio PHP e distribuito sotto la licenza GNU GPL.
Molti CMS contemporanei, compreso Drupal, offrono un'interfaccia di amministrazione, con cui l'amministratore può gestire ogni aspetto del sito web, senza il bisogno di possedere conoscenze tecniche particolari di programmazione

- Web: <<http://wiki.dublincore.org/index.php/Glossary>>
- FOAF** *Friend of a Friend Ontology.*
Un'ontologia comprensibile dai computer atta a descrivere persone, con le loro attività e le relazioni con altre persone e oggetti. Chiunque può usare FOAF per descriversi. FOAF permette a gruppi di persone di descrivere quel fenomeno noto come social network senza la necessità di un database centralizzato:
<<http://it.wikipedia.org/wiki/FOAF>>
- FRAD** *Functional Requirements for Authority Data.*
Un modello concettuale “Entità-Relazione” sviluppato dall’*International Federation of Library Associations and Institutions* (IFLA) che ha per oggetto di modellare i Record di autorità per autori e titoli. Record d’autorità (*Authority Record*) è la forma autorizzata del Nome combinata con altri elementi informativi che identificano e descrivono l’entità indicata e possono anche contenere rimandi ad altri record d’autorità collegati:
<http://www.ifla.org/files/assets/cataloguing/frad/frad_2013.pdf>
- FRBR** *Functional Requirements for Bibliographic Records.*
Uno schema concettuale sviluppato dall’IFLA , realizzato tramite modello “Entità-Relazione” allo scopo di dare una rappresentazione semi-formale alle informazioni bibliografiche:
<<http://www.ifla.org/en/publications/functional-requirements-for-bibliographic-records>>
- FRSAD** *Functional Requirements for Subject Authority Data.*
Uno schema concettuale sviluppato dall’IFLA , che ha per oggetto le registrazioni di autorità per i Soggetti:
<<http://www.ifla.org/files/assets/classification-and-indexing/functional-requirements-for-subject-authority-data/frsad-final-report.pdf>>
- ICCU** *Istituto Centrale per il Catalogo Unico delle Biblioteche Italiane e per le Informazioni Bibliografiche:*
<<http://www.iccu.sbn.it/opencms/opencms/it/>>
- IFLA** *International Federation of Library Associations and Institutions:*
<<http://www.ifla.org/>>
- IR** *Institutional Repository* o Archivio/Deposito/*Repository* Istituzionale
- ISAAR** *International Standard Archival Authority Record for Corporate Bodies, Persons and Families:* <http://www.icacds.org.uk/eng/isaar2ndedn-e_3_1.pdf>
- ISAR** *International Standard Archival Description:*
[http://www.icacds.org.uk/eng/ISAD\(G\).pdf](http://www.icacds.org.uk/eng/ISAD(G).pdf)
- KIF** *Knowledge Interchange Format.*
Un linguaggio di rappresentazione di conoscenza:
<<http://logic.stanford.edu/kif/>>
- KOSs** *Knowledge Organization Systems* o Sistemi di Organizzazione della Conoscenza (Liste di autorità), Sistemi di Classificazione, Thesauri, Topic Maps, Ontologies:
<http://en.wikipedia.org/wiki/Knowledge_Organization_Systems>
- LD** *Linked Data:* <<http://linkeddata.org/>>
- LOD** *Linked Open Data:* dati aperti, collegabili e usabili senza barriere:
<<http://www.linkedopendata.it/>>
- LDAP** *Lightweight Direct Access Protocol .*
Un Protocollo standard per l'interrogazione e la modifica dei servizi di directory come in generale qualsiasi raggruppamento di informazioni che può essere espresso come record di dati ed organizzato in modo gerarchico:

- LOM** <[http://it.wikipedia.org/wiki/Lightweight Directory Access Protocol](http://it.wikipedia.org/wiki/Lightweight_Directory_Access_Protocol)>
Learning Object Metadata.
Uno standard che fissa il minimo insieme di proprietà necessario per la gestione, l'allocazione e la valutazione dei *learning object*:
- METS** <[http://it.wikipedia.org/wiki/Learning Object Metadata](http://it.wikipedia.org/wiki/Learning_Object_Metadata)>
Metadata Encoding and Transmission Standard.
Uno Standard della *Library of Congress* per la codifica XML di metadati descrittivi, amministrativi, strutturali: <<http://www.loc.gov/standards/mets/>>
- MODS** <<http://www.loc.gov/standards/mods/>>
Metadata Object Description Schema.
Uno Standard della *Library of Congress* per la codifica XML di dati selezionati dai record di metadati MARC: <<http://www.loc.gov/standards/mods/>>
- Moodle** <<http://it.wikipedia.org/wiki/Moodle>>
Modular Object-Oriented Dynamic Learning Environment.
Un ambiente informatico *Open Source* per l'apprendimento modulare, dinamico, orientato ad oggetti (es. la gestione di corsi, basato sull'ideologia costruzionista secondo la quale ogni apprendimento sarebbe facilitato dalla produzione di oggetti tangibili): <<http://it.wikipedia.org/wiki/Moodle>>
- OAI** <<http://www.openarchives.org/>>
Open Archives Initiative. Un'organizzazione che produce le specifiche di interoperabilità per diverse piattaforme web: <<http://www.openarchives.org/>>
- OAI-PMH** <[http://it.wikipedia.org/wiki/Open Archives Initiative Protocol for Metadata Harvesting](http://it.wikipedia.org/wiki/Open_Archives_Initiative_Protocol_for_Metadata_Harvesting)>
Open Archives Initiative Protocol for Metadata Harvesting.
Un protocollo sviluppato dall'Open Archives Initiative come infrastruttura di comunicazione tra diversi metadati raccolti (o collezionati) dai documenti di Archivio – fornitori di dati (*Data Provider*) affinché possano essere costruiti i fornitori di servizi (*Service Provider*) utilizzando metadati da più *Data Provider*. Una implementazione dell'OAI-PMH deve supportare metadati rappresentati in Dublin Core, ma può supportare anche altre rappresentazioni:
<[http://it.wikipedia.org/wiki/Open Archives Initiative Protocol for Metadata Harvesting](http://it.wikipedia.org/wiki/Open_Archives_Initiative_Protocol_for_Metadata_Harvesting)>
- OAIS** <<http://public.ccsds.org/publications/archive/650x0m2.pdf>>
Open Archival Information System (Standard ISO 14721:2012).
Definisce concetti, modelli e funzionalità inerenti agli Archivi Digitali e agli aspetti di *Digital Preservation*:
<<http://public.ccsds.org/publications/archive/650x0m2.pdf>>
- OWL** <[http://it.wikipedia.org/wiki/Web Ontology Language](http://it.wikipedia.org/wiki/Web_Ontology_Language)>
Ontology Web Language.
Un linguaggio di *markup* di modellazione del Web Semantico per esprimere esplicitamente la semantica formale delle proprietà e Classi RDF:
<[http://it.wikipedia.org/wiki/Web Ontology Language](http://it.wikipedia.org/wiki/Web_Ontology_Language)>
- PREMIS** <http://www.loc.gov/standards/premis/Understanding-PREMIS_italian.pdf>
PREservation Metadata: Implementation Strategies.
Uno Standard di metadati di conservazione definiti come l'informazione usata da un Deposito/Archivio allo scopo di supportare il processo di conservazione digitale:
<http://www.loc.gov/standards/premis/Understanding-PREMIS_italian.pdf>
- RDA** <<http://www.rda-jsc.org/rda.html>>
Resource Description and Access.
Una specificazione basata sul modello concettuale FRBR che identifica un set di istruzioni per la catalogazione dei libri e di altro materiale bibliografico:
<<http://www.rda-jsc.org/rda.html>>
- RDF** <<http://www.w3.org/RDF/>>
Resource Description Framework.
RDF è una serie di specifiche sviluppate dal *World Wide Web Consortium* (W3C) per fornire uno strumento di modellazione dei metadati in una varietà di formati (sintassi):
<<http://www.w3.org/RDF/>>
- REST** <<http://www.html.it/guide/restful-web-services-la-guida/>>
Un Web Service RESTful è custode di un insieme di risorse sulle quali un agente web *client* può chiedere le operazioni canoniche sul protocollo web http. Propone una visione del Web incentrata sul concetto di risorsa:
<<http://www.html.it/guide/restful-web-services-la-guida/>>
- RIF** <<http://www.w3.org/2004/rif/>>
Rules Interchange Format. Una specifica del W3C per la descrizione di regole di

- interferenza per i metadati RDF:
<http://en.wikipedia.org/wiki/Rule_Interchange_Format>
- RSS** *Really Simple Syndication*. Una famiglia di formati XML usata per la pubblicazione dei contenuti frequentemente aggiornati:
<<http://it.wikipedia.org/wiki/RSS>>
- SKOS** *Simple Knowledge Organization System*.
Una specifica del W3C per rappresentare sistemi di organizzazione della conoscenza come Thesauri e Tassonomie usando i linguaggi RDF e OWL:
<<http://www.w3.org/2004/02/skos/>>
- SIP** *Submission Information Package*.
Pacchetto di Informazioni per l'immissione, utilizzato nella fase di immissione/acquisizione dei dati, mandato dal produttore in base al *Submission Agreement* stipulato con l'OAIS (*Open Archival Information System Standard*): <<http://www.iasa-web.org/tc04/submission-information-package-sip>>
- SOAP** *Simple Object Access Protocol*.
Un protocollo per lo scambio di messaggi tra componenti Software, tipicamente nella forma di componentistica Software, per mettere in risalto il concetto di servizio: <<http://it.wikipedia.org/wiki/SOAP>>
- SPARQL** *SPARQL Protocol and RDF Query Language*.
Una linguaggio di interrogazione del W3C per il modello di dati RDF:
<<http://www.w3.org/TR/rdf-sparql-query/>>
- SSD** Settori Scientifico-Disciplinari stabiliti dal Ministero dell'Istruzione, dell'Università e della Ricerca.
Sono raggruppamenti di discipline ai fini di agevolarne l'insegnamento a livello superiore nell'ordinamento italiano:
<http://it.wikipedia.org/wiki/Settore_scientifico-disciplinare>
- SSO** *Single Sign On*.
La proprietà di un sistema di controllo d'accesso che consente ad un utente di effettuare un'unica autenticazione valida per più sistemi software o risorse informatiche alle quali è abilitato: <http://it.wikipedia.org/wiki/Single_sign-on>
- SWAP** *Scholarly Works Application Profile*.
Un profilo applicativo Dublin Core per descrivere i lavori di ricerca:
<http://www.ukoln.ac.uk/repositories/digirep/index/Eprints_Application_Profile#Scholarly_Works_Application_Profile>
- TDR** *Trusted Digital Repository*:
<<http://www.oclc.org/content/dam/research/activities/trustedrep/repositories.pdf>>
- UML** *Unified Modeling Language*.
Un linguaggio di modellazione multi-scopo, grafico e orientato agli oggetti (*object-oriented*): <<http://www.uml.org/>>
- URI** *Uniform Resource Identifier*. È un riferimento universale e univoco per ogni risorsa disponibile sull'Internet.
The URI standard definition, RFC 2396: <<http://tools.ietf.org/html/rfc2396>>
- URL** *Uniform Resource Locator* si riferisce al sottoinsieme di URI che identifica le risorse attraverso una rappresentazione dei rispettivi meccanismi di accesso primario (per esempio la loro "locazione" nella rete), piuttosto che identificarle per nome o per mezzo di qualche loro attributo.
- VIAF** *Virtual International Authority File*.
Un progetto internazionale che ha dato vita ad un'unica base dati di voci di autorità controllate (ca. 20 milioni di autori provenienti da 24 cataloghi collettivi nazionali) a disposizione quale servizio gratuito del Web Semantico:
<http://it.wikipedia.org/wiki/Virtual_International_Authority_File>
- W3C** *World Wide Web Consortium*.

Un'organizzazione non governativa internazionale che ha come scopo quello di sviluppare tutte le potenzialità del World Wide Web:
<http://it.wikipedia.org/wiki/World_Wide_Web_Consortium>

XML *eXtensible Markup Language* . Una specifica del W3C per la codifica dei documenti in una forma letta dai computer: <<http://www.w3.org/XML/>>

XSL *eXtensible Stylesheet Language*.
Un linguaggio basato su XML per la trasformazione e la rappresentazione (*rendering*) di documenti XML: <<http://www.w3.org/Style/XSL/>>

INDICE

ABSTRACT	i
<i>ABSTRACT (en)</i>	iii
Struttura della Tesi	v
<i>Structure of the Thesis (en)</i>	vii
Ringraziamenti	ix
Acronimi	x
PREFAZIONE	1
Movimento internazionale <i>Open Access</i> e Archivi aperti Istituzionali di ricerca <i>Open Access</i>	1
Valore delle connessioni tra le unità informative abilitate come <i>Linked (Open) Data</i>	2
PARTE I	
<i>Open Access</i> e Archivi aperti per la comunicazione dell'informazione scientifica	6
CAPITOLO 1	
Movimento <i>Open Access</i>: origini, strategie, politiche e iniziative	11
1.0. Introduzione	11
1.1. Presupposti e dichiarazioni sull' <i>Open Access</i>	11
1.2. Stato dell'arte delle politiche <i>Open Access</i>	15
1.3. Strumenti, iniziative e attività a supporto del movimento <i>Open Access</i>	19
1.4. <i>Open Access</i> in Italia.....	22
1.5. Conclusioni	24
CAPITOLO 2	
Archivi aperti di ricerca <i>Open Access</i>: stato dell'arte, programmazione, questioni di etica e qualità	25
2.0. Introduzione	25
2.1. Progetto <i>Open Archive Initiative</i> a supporto delle piattaforme "Archivi aperti".....	25
2.2. Strategia dell' <i>Open Access</i> " <i>Green Road</i> ": Archivi aperti <i>Open Access</i>	28
2.2.1. Archivi aperti Istituzionali <i>Open Access</i> : scopi, funzionalità.....	31
2.2.1.1. Linee di indirizzo a supporto dell'implementazione e della gestione....	35
2.2.1.1.1. Raccomandazioni: il ruolo e alcuni casi d'uso.....	38
2.2.1.1.2. Linee guida: il ruolo e alcuni casi d'uso.....	39
2.2.1.2. Collegamenti con <i>Current Research Information Systems</i> (CRIS) per la valutazione della ricerca.....	41
2.2.1.2.1. <i>SurPlus Open Archive</i>	43
2.2.1.2.2. <i>U-GOV Ricerca</i>	46
2.3. Etica nel rispetto di contratti editoriali e proprietà intellettuale.....	50
2.4. Conservazione a lungo termine dei contenuti digitali.....	52
2.5. <i>Data Management and Process Plan (DMP)</i> : programmazione del ciclo di vita dei contenuti digitali.....	54
2.6. <i>Trusted Digital Repositories</i> : sostenibilità e affidabilità di Archivi digitali.....	56
2.7. Conclusioni	58

PARTE II	59
Indagine su alcuni strumenti a supporto degli Archivi digitali di ricerca.	
CAPITOLO 3	
Un'indagine su alcune piattaforme nazionali e internazionali a sostegno della comunicazione del sapere in rete e del Libero Accesso ai contenuti di ricerca	59
3.0. Introduzione.	59
3.1. Campionamento: selezione delle piattaforme	61
3.2. Domande d'indagine.	65
3.3. Schede compilative	67
3.4. Risultati dell'indagine.	91
3.5. Conclusioni. I fattori catalizzanti il valore degli Archivi Istituzionali	107
CAPITOLO 4	
ToolKit basato sullo stato dell'arte di alcuni aspetti a supporto del "Profilo Tecnico" di Archivi digitali di ricerca	111
4.0. Introduzione.	111
4.1. Panorama delle piattaforme.	114
4.2. Metadati a supporto dei contenuti digitali	124
4.3. Controllo di autorità dei valori di metadati.	133
4.4. Associazione di metadati con valori semantici.	142
4.5. Conservazione a lungo termine.	155
4.6. Digitalizzazione.	162
4.7. Scoperta delle risorse: paradigmi di ricerca e navigazione.	169
4.8. Conclusioni. La sintesi del contesto del "Profilo Tecnico" di un Archivio Istituzionale.	178
PARTE III	
Potenziamento dei metadati dalle strategie "LODE-BD" (<i>Linked Open Data – enabled Bibliographic Data</i>) per una migliore comunicazione dei contenuti di ricerca	187
CAPITOLO 5	
Sviluppo di uno schema di metadati a supporto descrittivo dei contenuti di ricerca negli Archivi Istituzionali alla luce delle strategie "LODE-BD" (<i>Linked Open Data – enabled Bibliographic Data</i>)	187
5.0. Introduzione.	187
5.1. Perché <i>Linked Data</i>.	189
5.2. Importanza dell'uso di vocabolari controllati e schemi semantici.	191
5.3. Alcuni approcci, documenti e principi per la creazione di metadati estensibili di qualità per descrivere i contenuti di ricerca.	199
5.3.1. Strategie delle "Linee guida per la creazione e la gestione di metadati nei Repository Istituzionali" (CRUI, 2012).	200
5.3.2. Strategie delle Raccomandazioni "LODE-BD" (<i>Linked Open Data – enabled Bibliographic Data</i>).	204
5.3.2.1. Alberi decisionali "Decisions Trees": scelta tra valori "letterali" e "non-letterali" per la codifica delle proprietà di metadati.	204
5.3.2.2. Alcuni vocabolari controllati per l'abilitazione dei metadati <i>RightsHolder, Licence, AccessRights</i> come <i>Linked Data</i>	207
5.3.2.3. Mappatura dei metadati con i marcatori di "<i>schema.org</i>" per una migliore comunicazione dei contenuti sul Web.	212

5.4.	Uno schema di metadati per la descrizione dei contenuti di ricerca negli Archivi Istituzionali basato sulle Linee Guida CRUI e le Raccomandazioni "LODE-BD".	212
5.4.1.	Strategie di implementazione dei metadati sviluppati alla luce di "LODE-BD".	236
5.5.	Conclusioni.	242
	<i>Conclusions (en)</i>	245
CONCLUSIONI della Tesi.		
	Comunicazione dell'informazione scientifica: le nuove frontiere	248
<i>CONCLUSIONS of the Thesis (en).</i>		
	<i>Communication of scientific information: new horizons.</i>	255
BIBLIOGRAFIA.		
SITOGRAFIA.		
LISTA DELLE FIGURE.		
LISTA DELLE TABELLE.		
APPENDICI		
Alcune modalità per la rappresentazione, gestione e l'interoperabilità dei contenuti digitali.		
317		
APPENDICE A		
Metadati e alcuni modelli a supporto della comunicazione dei contenuti digitali		
317		
1.1.	Strategie di metadati nella ricognizione di aspetti strutturali e semantici di contenuti digitali	318
1.1.1.	Metadati: funzioni e tipologie.	319
1.1.2.	Universo di Standard di metadati.	323
1.1.3.	<i>Dublin Core Metadata Element Set.</i>	325
1.2.	Web Semantico: nuovi approcci di comunicazione dei contenuti in rete.	328
1.2.1.	Linguaggi e soluzioni tecnologiche del Web Semantico.	330
1.2.1.1.	<i>Linked (Open) Data (Web 3.0): il valore delle relazioni per la posizione partecipativa dei contenuti sul web.</i>	334
1.2.1.2.	<i>Topic Maps: una tecnologia per organizzare e rappresentare la conoscenza.</i>	341
1.3.	<i>Unified Modeling Language (UML): un modello di classi e relazioni per lo sviluppo di Ontologie descrittive.</i>	344
1.4.	<i>Social Web o Web 2.0 per una cooperazione costruttiva dei contenuti digitali.</i>	349
APPENDICE B		
Alcuni approcci a supporto della programmazione e l'interoperabilità dei contenuti digitali di ricerca		
351		
2.1.	Principi Internazionali di Catalogazione dell' IFLA (<i>International Federation of Library Associations and Institutions</i>)	351
2.2.	Famiglia dei modelli concettuali FRBR (<i>Functional Requirements for Bibliographical Records</i>)	352

2.3.	<i>Dublin Core Application Profile (DCAP): creazione di profili di metadati applicativi per domini specifici.</i>	353
2.3.1.	Profilo applicativo per i contenuti di ricerca <i>Scholarly Works Application Profile (SWAP)</i>	355
2.4.	Schema di metadati proposto dal Gruppo <i>Open Access</i> (CRUI, Italia, 2012)	357
2.5.	<i>Core Scientific Meta-Data Model (CSMD)</i>	362
2.6.	Interoperabilità, integrazione, condivisione dei contenuti di ricerca.	359
2.6.1.	Rapporto “ <i>The Current State of Open Access Repository Interoperability</i> ”	360
2.6.2.	Standard <i>Common European Research Information Format (CERIF)</i> a supporto dell’interoperabilità tra vari sistemi informativi di ricerca.	361
2.6.3.	<i>VIVO Ontology</i> : la rappresentazione di entità collaborative coinvolte nei processi di ricerca.	362
2.6.4.	<i>Citation Typing Ontology (CITO)</i>	363
2.6.5.	<i>Extensible Observation Ontology(OBOE)</i>	364
2.6.6.	Nano-pubblicazioni: un approccio per comunicare le entità concettuali di un prodotto di ricerca.	365
2.6.7.	<i>Open Archives Initiative – Object Reuse and Exchange Protocol</i> a supporto del modello “ <i>Enhanced Publication</i> ”	367
2.6.8.	Approcci di aggregazione di oggetti digitali complessi nella Biblioteca digitale europea <i>Europeana</i>	373
2.6.9.	Raccomandazioni “ <i>LODE-BD</i> ” (<i>Linked Open Data – enabled Bibliographic Data</i>)	370
2.7.	Progetto dell’Archivio di ricerca <i>Open Access VOA3R</i>	373
2.7.1.	Programmazione dei requisiti di metadati estensibili di qualità	373
2.7.2.	Costruzione del sociale (Web 2.0) e semantico (<i>Linked Data</i>) per la comunicazione dei contenuti digitali.	376
2.7.2.1.	Funzionalità avanzate di ricerca e di valutazione della qualità di contenuti di ricerca.	382

PREFAZIONE

Movimento internazionale *Open Access* e Archivi aperti Istituzionali di ricerca *Open Access*

L'*Open Access* [Suber, 2012] è un movimento internazionale per promuovere la disseminazione del sapere scientifico e il libero accesso ai contenuti di ricerca, che si focalizza maggiormente su quelli finanziati con i fondi pubblici su vasta scala, attraverso politiche e strategie *Open Access*, la rete Internet e le tecnologie informatiche all'avanguardia per la gestione e diffusione dei contenuti digitali.

La *Budapest Open Access Initiative* (2002) – il primo raduno storico di fondazione dell'*Open Access* - ha formalizzato la nascita di due strategie (canali) per la disseminazione alternativa dei contenuti di ricerca:

1) “*Green Road*” (“via verde”) – gli Archivi aperti *Open Access* (Disciplinari e Istituzionali) per l'auto-archiviazione (deposito) dei materiali di ricerca (*pre-prints, post-prints*);

2) “*Gold Road*” (“via dell'oro”) - la pubblicazione di articoli nelle Riviste *Open Access*.

La definizione del materiale di ricerca ad Accesso Aperto si associa al concetto “*BBB Definition*” basato sulle definizioni promosse da:

a) *Budapest Open Access Initiative* (2002);

b) *Bethesda Statement on Open Access Publishing* (2003);

c) *Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities* (2003):

La letteratura ad Accesso Aperto (*Open Access, OA*) è *digitale, online, gratuita e libera da buona parte delle restrizioni* dettate dalle Licenze per i diritti di sfruttamento commerciale. Queste condizioni sono possibili grazie ad Internet e al consenso dell'autore o del titolare dei diritti d'autore [...] Gli scienziati e gli studiosi sono diversi dai musicisti e dai produttori di film: le controversie relative all'Accesso Aperto alla *musica* e alle *pellicole cinematografiche* non si applicano alla letteratura di ricerca [...] La letteratura ad Accesso Aperto ha dei costi di produzione, sebbene inferiori a quelli delle pubblicazioni commerciali. La questione non riguarda come abbassarli, ma se vi siano dei modi di coprirli che non ricadano sui lettori e non creino barriere all'Accesso. I modelli economici dell'Accesso aperto dipendono dalle modalità con cui lo si fornisce [V. sito “Breve Introduzione all'Accesso Aperto”].

L'*Open Access* è al centro di numerosi dibattiti, attraverso i quali Università, ricercatori, bibliotecari e politici si confrontano sui concetti e le implicazioni economiche, politiche e scientifiche che ne derivano.

Gli Archivi aperti Istituzionali *Open Access*, sviluppati sulle piattaforme tecnologiche *Open Archives* e appoggiati sulle politiche *Open Access*, offrono i luoghi virtuali per raccogliere, documentare, conservare e disseminare la produzione dell'attività scientifica svolta dagli studiosi nell'ambito della loro attività istituzionale.

Un Archivio Istituzionale viene paragonato ad “*una vetrina della produzione scientifica dell'Ente, un modello di comunicazione e disseminazione della conoscenza nell'era di Internet, ovvero un servizio d'informazione e documentazione rivolto alla collettività*” [Guerrini, 2010, p.28], che aumenta la visibilità, la citazionabilità, il prestigio e il valore pubblico della ricerca di un Ente a livello globale tramite la rete.

Valore delle connessioni tra le unità informative abilitate come *Linked (Open) Data*

*The challenge is not to make data
theoretically accessible, but also practically usable*
[G-8 International Conference on Open Data, 2013]

Il concetto *Linked Data* (o Web di Dati collegati, *Web 3.0*) - coniato nel 2006 dall'inventore del Web Tim Berners Lee (Berners Lee, 2006) - fa riferimento ad una serie di "*Best Practices*" (sito '*Best Practices for Publishing Linked Data*') per strutturare, connettere e pubblicare *dataset* (un insieme di dati) secondo i linguaggi e metodi del Web Semantico.

L'idea di Web Semantico può essere affrontata anzitutto come un *fenomeno linguistico*. Il metodo che permette l'integrazione di dati diversi in un modo coerente si presenta come una "lingua" per i dati [Si tratta] di una lingua inventata dagli umani, una *lingua inventata per comunicare informazioni* e pensieri fondamentalmente umani e per l'uso umano, una lingua fatta però per essere *letta ed elaborata* non da lettori o ascoltatori umani bensì *da computer* secondo algoritmi ben definiti [Baker, 2012, p.1].

La tecnologia *Linked Data* ha l'obiettivo di estendere il potenziale del Web attuale (o Web di documenti, Web ipertestuale) dove la semantica del contenuto e i legami tra i documenti sono strutturati dagli umani, ma non possono essere interpretati dalle macchine.

Con *Linked Data* possono essere strutturate e formalizzate in modo articolato diverse unità informative (oggetti di informazione come ad es. dati, metadati, pagine, documenti, collezioni, voci di vocabolari, nomi personali e geografici) disponibili in formato digitale, dando luogo alla creazione di una *rete* di relazioni *radicate* nella semantica delle unità informative stesse.

Il presupposto importante per abilitare e pubblicare diverse tipologie di unità informative come *Linked Data* [Villazón-Terrazas et al., 2011] è avere a disposizione per la loro identificazione gli indirizzi univoci web *Uniform Resource Identifiers* (URIs) appoggiati sul protocollo internet HTTP

(es. <<http://leo.cilea.it/index.php/jlis/article/view/6307/7884>>, <<http://dewey.info/scheme/e22/>>) per essere individuati sul Web da persone, browser e programmi. Il secondo presupposto è organizzare gli identificatori URI_HTTP entro le stringhe di un linguaggio per la modellazione del Web Semantico - *Resource Description Framework* (RDF). RDF è una sintassi leggibile e interpretabile sia dagli esseri umani che dalle macchine ed è costituita da una tripla ovvero dai tre segmenti collegati tra loro:

- 1) "*Soggetto*" (l'identificatore di un'unità informativa - *Risorsa*);
- 2) "*Predicato*" (l'identificatore di un'unità informativa che definisce una certa proprietà del "Soggetto" – tramite una *Relazione*);
- 3) "*Oggetto*" (l'identificatore di un'unità informativa che descrive un *Valore* della *Relazione*).

I *Linked Data* costituiscono un *nuovo linguaggio* nel mondo della *comunicazione globale* [...] i dati sono collegati ad altri set di dati esterni, a cui altri set di dati possono a loro volta riferirsi, creando un reticolo sempre più vasto, costituito da Ontologie, cioè da *raggruppamenti logicamente definiti* [Guerrini, 2013, Introduzione],

attraverso i quali possono essere sviluppate numerose connessioni tra unità informative (*dataset*) create da diversi soggetti produttori, che possono

includere link ad altri URI relativi ai dati esposti per migliorare la ricerca nel Web di altre informazioni affini a quella di partenza: più i dati sono collegati, più sarà possibile il loro utilizzo nell'ottica di arricchimento e deduzione delle informazioni [Guerrini, Posemato, 2012, p.12].

Il panorama di unità informative pubblicate e interconnesse per mezzo di *Linked Data* viene rappresentato nella *Linked Open Data Cloud* (Figura 1), la cosiddetta "nuvola" di *dataset* pubblicati con la tecnologia *Linked Data*. La maggior parte di *dataset* sono pubblicati con

Licenze Aperte e in formati aperti [V. sito "*Open Data Charter and Technical Annex*"; De Robbio, 2012].

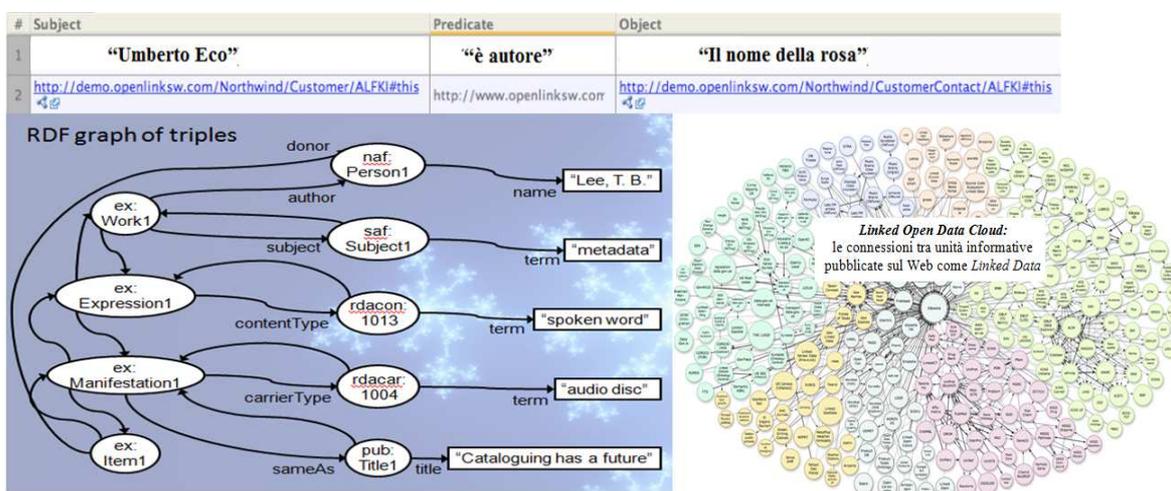


Figura 1. Una rappresentazione grafica dei collegamenti RDF tra le unità informative di un Catalogo e tra dataset nella *Linked Open Data Cloud* [Dunsire, 2012]

La strutturazione dei nodi interconnessi del modello RDF “Soggetto” (es. “Umberto Eco”) - “Predicato” (es. “è autore”) – “Oggetto” (es. “Il nome della rosa”) può essere paragonata alla struttura di database relazionali costituita tramite tabelle realizzate da righe (*record*) e colonne, sul cui incrocio si ottiene una serie di campi, i quali possono rispecchiare la rappresentazione dei nodi RDF. Mediante le relazioni costituite dal modello RDF si può creare una varietà di ontologie intese come reti relazionali tra diverse asserzioni (elementi, atomi) di unità informative.

La forza dei *Linked Data* sta nel fatto che ogni asserzione della tripla RDF diventa - a sua volta - generatrice di nuove triple e, quindi, di nuove unità informative, agganciate tra loro mediante gli URIs (*Uniform Resource Identifiers*) basati sul protocollo internet HTTP - creando reticoli (grafi) di triple sempre più estese, grazie al principio dell’inferenza possibilitata dai meccanismi del Web Semantico. Le unità informative abilitate come *Linked Data* diventano le parti di collezioni di conoscenza sempre più ampie e significative, leggibili e ri-utilizzabili dalle piattaforme digitali Web.

Nell’ottica delle unità informative pubblicate come *Linked Open Data*, tre sono i fattori di maggiore interesse:

- 1) la semantica dei dati;
- 2) la loro libera o aperta disponibilità per l’utente in rete:

Nella scienza è fortemente raccomandato che i dati siano esplicitamente di dominio pubblico, soprattutto quando sono finanziati con denaro pubblico. I dati dovrebbero essere quindi rilasciati con una Licenza di tipo *Public Domain Dedication Licence (PDDL)* oppure *Creative Commons Zero Waiver*, o *CC0* [De Robbio, 2012, p.12];
- 3) la fruibilità e il riuso di atomi di unità informative in una molteplicità di soluzioni di collegamenti.

La tecnologia *Linked Data* potrebbe essere molto utile per strutturare e pubblicare una varietà delle rappresentazioni di unità informative ospitate entro le piattaforme di Biblioteche e Archivi digitali, cui contenuti spesso sono difficilmente trovabili sul Web. Considerando che sempre più utenti:

- a) tendono ad iniziare una ricerca di informazioni utilizzando un motore sul Web piuttosto che il tradizionale Catalogo o una piattaforma dell’Archivio digitale;
- b) gradiscono rilevare sul Web i contenuti correlati nel significato,

le soluzioni *Linked Data* possono presentare un punto di partenza strategico per abilitare, integrare e collegare le unità informative di Biblioteche e Archivi entro lo spazio del Web.

Produrre *Linked Data* significa, dunque, esprimere i *significati* delle informazioni, renderle *condivisibili* fra differenti applicazioni e *utilizzabili* da applicazioni diverse da quelle per cui erano state originariamente create

[De Robbio, 2012, p.12].

Per i possessori di grandi quantità di dati la tecnologia *Linked Data* rappresenta un'innovazione suscettibile per lo sviluppo di una diversità di applicazioni a valore aggiunto. Un esempio pratico è il servizio di ricerca dei contenuti digitali offerto dalla piattaforma "*OpenAGRIS*" il quale rivela non solo le risorse informative ospitate entro la stessa piattaforma, ma anche le risorse correlate provenienti dalle piattaforme esterne. Questo tipo di "accesso multifocale" ai contenuti correlati nella semantica è possibile grazie alle relazioni create a partire dall'impiego negli schemi di metadati (indici descrittivi di risorse informative) dei valori del Thesaurus AGROVOC disponibili online in modalità *Linked Data* (Figura 2).

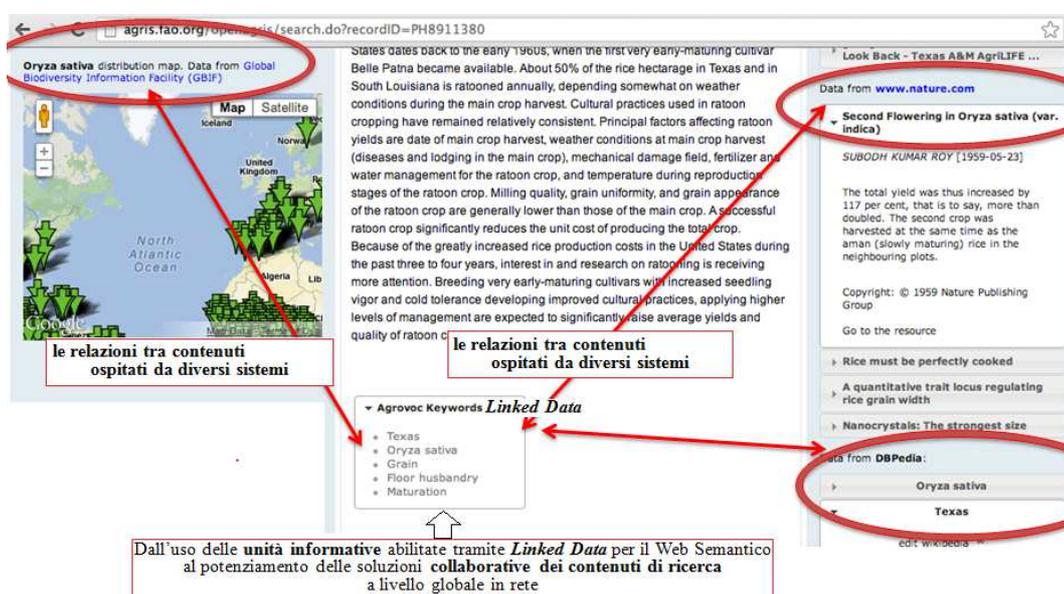


Figura 2. La piattaforma di ricerca "*OpenAGRIS*" [V. sito] che impiega le potenzialità del Thesaurus AGROVOC in modalità *Linked Data*

Un soggetto produttore dei contenuti digitali - che intende di abilitare le proprie fonti informative (es. metadati, documenti, vocabolari controllati) come *Linked Data* - può ricorrere alle seguenti soluzioni:

- 1) pubblicare direttamente le proprie fonti informative come *Linked Data* [v. sito: la *Classificazione Decimale Dewey* pubblicata come *Linked Data*] e registrarle sul *CKAN Registry* (*Comprehensive Knowledge Archive Network*; v. sito) - per rendere possibile la loro scoperta, la condivisione e il riutilizzo;
- 2) impiegare i *dataset* disponibili come *Linked Data* (es. l'uso delle voci del Thesaurus *Linked Data* AGROVOC nel servizio "*OpenAGRIS*") per qualificare i valori di metadati descrittivi le fonti informative;
- 3) (1) e (2) assieme.

In ambedue i casi è importante poter operare la formalizzazione e il collegamento dei dati tramite le piattaforme *TripleStores*, ovvero, database per ospitare e gestire le triple RDF, con cui modellare le relazioni tra unità informative in modo da definire un contesto semantico desiderato e funzionale (Figura 3). Diversi *Software Open Source* (es. *Fedora*, *Dspace*, *E-Prints*) preposti

per l'implementazione delle piattaforme di Archivi aperti sono i *Software TripleStores* ovvero i *Software* semantici.

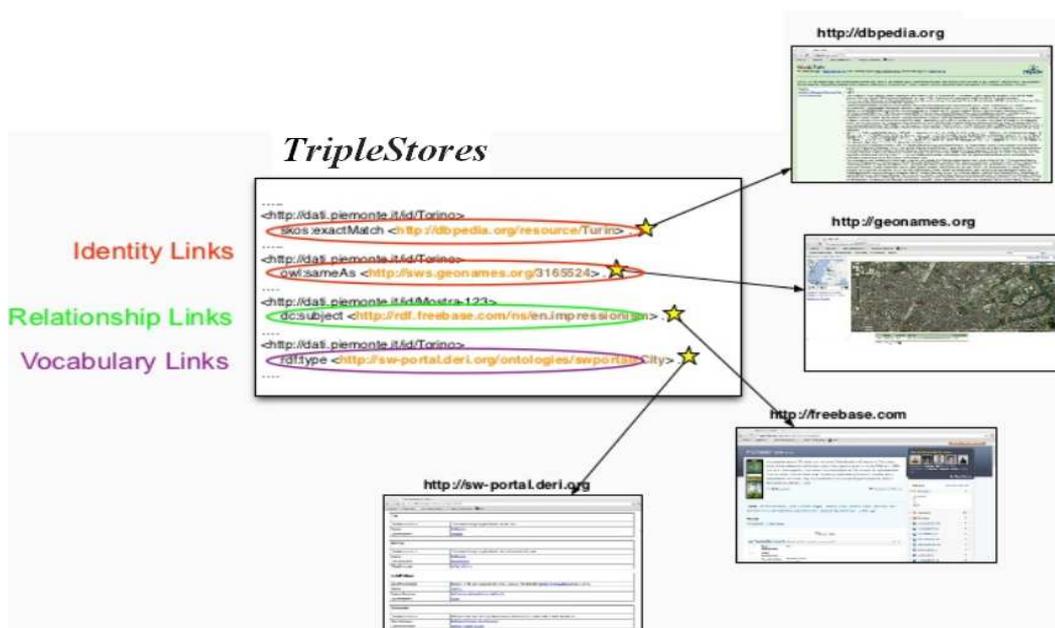


Figura 3. La formalizzazione di unità informative e di loro collegamenti entro le piattaforme *TripleStores*

Quegli Archivi che sapranno combinare l'apertura e la pubblicazione di loro unità informative tramite la tecnologia *Linked Data*, potranno ritagliarsi "spazi di prestigio" nei reticoli e aggregazioni di *Linked Open Data Cloud* del Web Semantico, facendo sì che le medesime unità potranno diventare riusabili (ricollegabili, ricomponibili e ricombinabili) da diversi soggetti partecipanti nella produzione e pubblicazione dei contenuti (dati) digitali sul Web.

I dati rappresentano la *materia necessaria* per creare l'ecosistema ma, per ottenere l'attivazione di questi processi, occorrono dei *campioni* (intesi come attori del quadro economico, *istituzionale* o sociale che riescono a diventare esempi da seguire ed emulare) che interpretino al meglio quanto oggi è presente in termini di tecnologia, norme e comunicazione

[V. sito: Linee Guida per l'interoperabilità semantica attraverso i *Linked Open Data*, 2012, p. 81]

Per i campioni di unità informative dell'ecosistema "Archivi Istituzionali" [V. sito: "What are Institutional Repositories"] la tecnologia *Linked Data* può presentare una forte leva d'innovazione, una "formula mista", che non affida la selezione delle unità informative solo alle macchine (tecnologia semantica) tramite mediazione di "parole chiave", ma propone la formula per la creazione del dialogo tra diverse unità informative, rendendo possibile così un dialogo interattivo "uomo-macchina" di reciproco arricchimento.

Tramite *Linked Data* possono essere:

- a) erogati i servizi di ricerca semanticamente ricchi;
- b) aumentate le citazioni al materiale di ricerca depositato negli Archivi Istituzionali, a partire dalle relazioni create tra diverse unità informative digitali, semanticamente correlati.

Il progetto "*dati.camera.it*" (Italia) fornisce un buon esempio del servizio di ricerca dei contenuti digitali (pubblicati come *Linked Open Data*) correlati tramite i navigatori RDF:

Per visualizzare e consultare i dati è possibile utilizzare dei navigatori RDF [che offrono] un *quadro esaustivo della risorsa* in esame esplicitando tutti i collegamenti con le altre informazioni (fotografie, documenti pdf, ecc.) [V. sito: "*dati.camera.it*"].

PARTE I

Open Access e Archivi aperti per la comunicazione dell'informazione scientifica

È la natura delle tecnologie digitali a reclamare l'apertura [...] per rendere il mondo della scienza più equo, democratico e innovativo [Caso, Galimberti, 2013]

With today's ability to share knowledge over the web, it just *doesn't make sense* that *so much research is locked away* from the majority of the world [Open Access Button, 2013, V. sito]

Per le Università poter disporre di una base estesa di pubblicazioni e dati in Accesso Aperto significa garantirsi libero e gratuito accesso al *materiale di base della ricerca* (e della didattica), materiale altrimenti soggetto a regole fortemente restrittive [Caso, 2013, p.50]

Lo scopo è rimuovere le barriere *d'accesso alla letteratura scientifica*, favorendone la *più ampia disseminazione e condivisione*, e, ancor più, migliorare le procedure di *valutazione qualitativa*, in un contesto di riferimento che è sempre meno ancorato a realtà locali o regionali [Guerrini, Crupi, 2012, p.3]

I soggetti pubblici preposti all'erogazione o alla gestione dei finanziamenti della ricerca scientifica adottano [...] le misure necessarie per la *promozione dell'Accesso Aperto ai risultati della ricerca finanziata* per una quota pari o superiore al 50 per cento con *fondi pubblici* [V. Legge n. 112 del 7 ottobre 2013, la normativa sull'Accesso Aperto]

Per comunicazione dell'informazione scientifica

si intendono i vari modi in cui avviene lo scambio di informazioni tra gli studiosi: include quindi la *produzione e diffusione delle pubblicazioni* delle Istituzioni di ricerca, la *comunicazione informale* tra studiosi [...] la *discussione* ai convegni e ai seminari, l'*accesso* all'informazione attraverso la ricerca in cataloghi di biblioteche, banche dati e in altri *strumenti di comunicazione in rete*. In un ambito digitale tutte queste funzioni – *le pubblicazioni* e *la comunicazione* – diventano una sola: la *diffusione dell'editoria scientifica* che si basa sull'*interattività* e *dinamicità* del documento digitale [Tammaro, 2001, p.22-34].

Nonostante l'Internet e le nuove tecnologie digitali configurino nuovi modi di pensare e nuove dinamiche sociali che dovrebbero allargare gli orizzonti e abbattere significativamente le barriere della comunicazione dell'informazione scientifica, si continua ad assistere ad una serie di paradossi.

Nel corso di un convegno nel 1992, Joshua Lederberg, della *Rockefeller University*, premio Nobel per la medicina nel 1958, denunciò il crescente divario tra *i costi delle Riviste accademiche* e *gli stanziamenti destinati alle Biblioteche* per il loro acquisto. Era entrato in *crisi* il *modello tradizionale di editoria scientifica*, che consisteva in una rete di *Riviste specializzate* gestite da un numero ristretto di editori e vendute a prezzi ritenuti cari e talora esosi [Guerrini, 2010, p.13].

La rapida e incontrollabile crescita dei costi di abbonamento per le riviste scientifiche (“*serials pricing crisis*” [Hall, 2009]) ha portato molte Biblioteche all’impossibilità di garantire una *crescita delle collezioni* proporzionata alla *crescita della ricerca scientifica*.

I costi elevati delle Riviste condizionano fortemente la loro diffusione e ne restringono la circolazione alle istituzioni più ricche (con l’aumento vertiginoso delle spese sostenute dalle Università), privandone quelle meno capienti, che si vedono costrette ad operare tagli massicci di risorse in abbonamento con una pari limitazione dell’offerta informativa indispensabile alla ricerca.

Nel documento “*Policy Perspectives*” dell’*Association of Research Libraries (ARL)* viene constatato il fatto che

commercial publishers have made up in price what they have lost in volume, having understood that where demand for their product is strong, the pressure on Libraries to maintain currency in a field will continue regardless of price [V. sito: *Policy Perspectives*].

Nonostante i tagli agli abbonamenti alcuni editori registrano ancora profitti molto alti, in alcuni casi pari al 40%. Il fenomeno “*serials pricing crisis*” è soprattutto svantaggioso per la disseminazione delle informazioni scientifiche nelle istituzioni dei paesi in via di sviluppo nei quali ci sono diverse Università non abbonate proprio alle Riviste scientifiche [V. sito: Rete cooperativa]. In questo scenario gli autori spesso sono costretti ad avere l’accesso solo ad una parte della letteratura di loro interesse offerta dalla loro istituzione.

Con questo quadro di riferimento, anche la promessa dell’*online* si è rivelata poco promettente per i risultati della ricerca, in quanto i *costi per l’accesso* (“*permission crisis*”) alle pubblicazioni elettroniche che devono sostenere le Biblioteche accademiche sono *sempre più elevati* e sono, per lo più, in *aggiunta a quelli per il cartaceo*.

Le barriere legali derivano dalle leggi sul diritto d’autore e dai contratti di edizione, spesso assai restrittivi; le barriere tecnologiche dall’uso di sistemi di *Digital Rights Management (DRM)*, *software* in grado di bloccare l’accesso a utenti non autorizzati [Caso, 2006].

Molte istituzioni accademiche continuano a sostenere il costo della ricerca, ne cedono gratuitamente agli editori i risultati e infine li riacquistano a costi ormai quasi intollerabili:

Il rincaro del costo degli abbonamenti penalizza il sistema delle Biblioteche delle Università [...] che ricoprono [...] la duplice veste di finanziatori del lavoro dei propri ricercatori e di acquirenti delle riviste sulle quali i ricercatori pubblicano la ricerca finanziata dalla stessa Università [Segnini, Vespucci, 2004].

“Il risultato è paradossale: gli Atenei sostengono il costo della ricerca, ne cedono gratuitamente agli editori i risultati e infine li riacquistano a costi ormai quasi intollerabili” [V. sito: *Open Access e comunicazione scientifica*] (Figura 4).

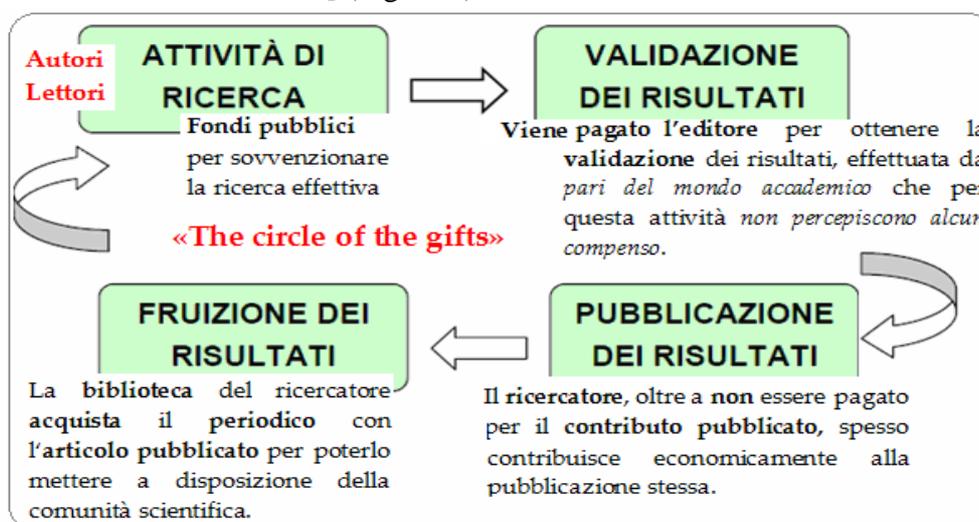


Figura 4. “*The circle of the gifts*” [Case, 1999] (regalo di scambio): il paradosso della proprietà intellettuale nella comunicazione dell’informazione scientifica

Diverse biblioteche delle istituzioni accademiche di tutto il mondo non riescono quindi a tenere il passo con la mole e i costi sempre più elevati delle riviste scientifiche e ciò ha causato continue cancellazioni di abbonamenti da parte di diverse biblioteche. Come effetto: gli autori comunicano esclusivamente con coloro che afferiscono ad un’istituzione in grado di pagare l’accesso alle loro pubblicazioni ed i lettori accedono, per conseguenza, solo ad una parte della letteratura di loro interesse.

Solitamente gli accademici sono pagati per le loro ricerche o dalle Università o comunque con finanziamenti pubblici; l’articolo pubblicato è il risultato del loro lavoro, non un mezzo per guadagnare denaro [De Robbio, 2007, p.53].

I ritardi nella pubblicazione degli esiti delle ricerche spesso rallentano la diffusione del materiale di ricerca all’interno della comunità scientifica. I tempi di latenza tra l’invio di un articolo e la sua pubblicazione sono normalmente misurati in mesi, sempre che il contributo venga poi effettivamente accettato. Intanto, più immediatamente un articolo è letto, più è citato. Di conseguenza, è meglio sia per la ricerca, sia per la carriera del ricercatore.

La crisi che ha toccato il modello tradizionale della comunicazione scientifica, come ogni crisi, costituisce anche un’opportunità di cambiamento. Infatti, molte istituzioni accademiche, subendo i fattori negativi che ostacolano la comunicazione della loro informazione scientifica, hanno in realtà generato al loro interno un fermento di iniziative e progetti senza precedenti, destinati a cambiare profondamente le modalità di circolazione del sapere.

A seguire, sono fornite solo alcune tra numerose iniziative introdotte da diverse istituzioni per la disseminazione libera dei contenuti di ricerca in formato digitale:

- 1) la creazione negli anni 90 dello scorso secolo degli Archivi digitali per la disseminazione libera dei risultati di ricerca nei campi di fisica ed economia “ArXiv” e “RePec”;
- 2) la messa in atto all’inizio del nuovo secolo delle nuove forme dell’editoria digitale entro diversi Atenei come, ad esempio, “*Firenze University Press*” [v. sito; Guerrini, Ventura, 2009], “*Cambridge University Press*” [V. sito], gli strumenti privilegiati di una nuova attività editoriale che hanno inteso esternalizzare dalle restrizioni introdotte dalla editoria commerciale, con l’obiettivo che la comunicazione scientifica tornasse a circolare liberamente tra coloro che ne sono, al contempo, produttori e fruitori;
- 3) l’istituzione della casa editrice *Open Book Publisher* [V. sito] all’Università di Cambridge che, appoggiandosi a *GoogleBooks*, offre la lettura gratuita dei libri prodotte entro la stessa Università;
- 4) la costituzione del *Network* nazionale di Archivi Digitali Istituzionali in Argentina [V. sito] che, promuovendo le *Policy* condivise, permette alle Istituzioni, che ne fanno parte, di coordinare piani d’azione e definire standard generali nella costruzione di un’infrastruttura digitale globale basata sugli Archivi digitali di Ricerca visibili su scala internazionale. Lo scopo principale dell’iniziativa è promuovere l’Accesso Aperto, favorire lo scambio nella ricerca attraverso la rete nazionale degli Archivi e l’interoperabilità degli stessi con altri network sudamericani e nel mondo;
- 5) la decisione di rendere le riviste dell’Archivio Storico della *Royal Society* [V. sito] (Accademia Nazionale di Scienze britannica) ad Accesso Aperto in modo permanente. Si tratta di più di 60000 articoli scientifici pubblicati a partire da oltre 70 anni fa. L’iniziativa si inserisce nella strategia della *Royal Society* in direzione dell’editoria ad Accesso Aperto che è “contro i rischi di privatizzazione o di mero sfruttamento economico dell’informazione pubblica” [V. sito AIB-DFP].

L’Accesso Aperto (*Open Access, OA*) è una serie di iniziative e strategie internazionali che promuovono il *modello alternativo di circolazione della conoscenza scientifica*, considerato

idoneo a risolvere le anomalie e le contraddizioni che caratterizza spesso l'attuale sistema monopolizzato dall'editoria commerciale e fondato su presupposti culturalmente ed economicamente poco vantaggiosi [Pievatolo, 2013].

L'Accesso Aperto alla comunicazione dell'informazione scientifica ha l'obiettivo di promuovere:

1. l'accesso libero e senza barriere al *sapere scientifico finanziato con fondi pubblici* [V. sito: *guardian.co.uk*], che si traduce in significativi benefici economici, sociali ed educativi, in quanto i risultati delle ricerche finanziati con denaro pubblico devono essere di dominio pubblico [Melanie Dulong de Rosnay, Juan Carlos De Martin, 2006; V. sito "Public access to publicly-funded research"]: "Every citizen should have Free Open Access to publicly funded research, data and knowledge" [V. sito: *Brisbane Declaration*];
2. la progettazione di *nuovi modelli editoriali* per dare rilievo al valore della produzione interna degli Atenei, per accelerare la circolazione dei risultati della ricerca di qualità (con metriche di valutazione adeguate, eque e trasparenti) gratuitamente e senza restrizioni d'accesso [V. sito: *Research LibraryUK/SCONUL*], affinché essi raggiungono l'impatto [V. sito: *Special issue on the state of the art in assessing research impact*] utile ad uno sviluppo collettivo, proponendo anche nuove forme di trattamento del *Diritto d'Autore*: "il passaggio all'Accesso Aperto modifica la disseminazione della conoscenza nei suoi aspetti organizzativi, legali e finanziari [V. sito: "Dichiarazione di Berlino"];
3. *dati (contenuti) aperti* della ricerca pubblicati sotto le *Licenze aperte*.

Perché la scienza si costruisce con l'accumulo delle esperienze, poter attingere ai dati di ricerche o esperimenti incoraggia questo processo.

Per lo più, l'*Open Access* è la garanzia più efficace *contro i rischi di plagio*, poiché aumenta a dismisura le possibilità che il plagio venga scoperto evitando così le duplicazioni di ricerche. La qualità degli Atenei a livello internazionale si misura oggi sul grado di eccellenza della didattica e sull'originalità e il valore dei risultati della ricerca. Avere a disposizione strumenti automatizzati, che facilitano la rilevazione del plagio nei prodotti della didattica e della ricerca, appare strategico, non tanto per l'attuazione di politiche punitive, ma proprio per l'avvio di processi di continuo miglioramento.

Accanto alla strategia "*Gold Road*" (via d'oro)¹ (Riviste ad Accesso Aperto), vi è la strategia parallela "*Green Road*" (via verde), che promuove la creazione di Archivi Istituzionali (e anche Disciplinari) digitali di ricerca *Open Access*. Attraverso gli Archivi Istituzionali è possibile promuovere online la conoscenza sulla produzione scientifica degli Atenei ed Enti di Ricerca a livello globale, valorizzando maggiormente il materiale scientifico pubblicato in modalità *Open Access*. Ciò prevede la consultazione e l'uso del materiale da parte degli utenti in rispetto alle appropriate Licenze d'uso.

Nel prossimo futuro sarà un'anomalia per una Università o per un Ente di ricerca non avere il proprio Archivio Istituzionale: non ci si può privare di uno strumento unico e strategico per pubblicizzare la produzione intellettuale dell'Istituzione, massimizzandone la visibilità e l'impatto nei confronti dei vari portatori di interesse

[V. sito "Linee guida per gli Archivi Istituzionali", introduzione].

Nel 2007 la Comunicazione della Commissione al Parlamento Europeo [V. sito] ha lanciato il seguente messaggio: "Sono indispensabili iniziative che favoriscano un Accesso più ampio e una

¹ Da non confondere con la "Via Rossa" (*Red Road*), che prevede il "*double dipping*" riconducibile al pagamento da parte di autori di una quota, una *fee* (es. *Article Processing Charge*, APC, conosciuta dalle decisioni del Governo Willets britannico) a sostegno delle spese di pubblicazione di loro articoli nelle riviste chiuse, in modalità *Toll-Access*, che non rispecchia il modello *Open Access* (OA). La prassi APC, intanto, via via si sta modificando inglobandosi nella modalità APC verso l'OA.

maggior diffusione dell'informazione scientifica, soprattutto per quanto concerne *Articoli di riviste e Dati di Ricerca* prodotti con *Finanziamenti Pubblici*".

Nel 2012 - a dieci anni dal primo raduno storico sull'*Open Access*, la *Budapest Open Access Initiative* (BOAI, 2002) in cui sono stati formalizzati i principi dell'*Open Access* - è stato stilato il Programma per i prossimi dieci anni di attività per la promozione dell'Accesso Aperto al materiale scientifico [V. sito "Programma BOAI per i prossimi dieci anni"].

Tra i punti principali di tale programma ci sono:

1. La diffusione delle Politiche *Open Access* all'interno degli Atenei. Tali Politiche devono:

a) garantire il deposito (auto-archiviazione) da parte dei ricercatori di una copia (nella versione *peer-reviewed*) della loro produzione scientifica (articoli, set di dati di ricerca, presentazioni a convegni, monografie o capitoli di monografie) negli Archivi Istituzionale preposto;

b) incoraggiare la pubblicazione di articoli in Riviste *Open Access* ma non renderlo obbligatoria, facendo inoltre comprendere la differenza tra "depositare in un Archivio *Open Access*" e "pubblicare su una Rivista ad Accesso Aperto".

2. L'applicazione della Licenza CC-BY per le Rivista ad Accesso Aperto e per il materiale da depositare negli Archivi (Istituzionali e Disciplinari), quando sia possibile farlo nel rispetto dei diritti acquisiti da autori ed editori.

3. La sostenibilità delle infrastrutture tecnologiche di Archivi (*Repository*) aperti che devono:

a) consentire l'*harvesting* (la raccolta) di metadati (a supporto descrittivo di contenuti) da altri Archivi e il ri-deposito di contenuti in altri Archivi;

b) sviluppare statistiche d'uso, citazioni e altri servizi a valore aggiunto, al fine di favorire la creazione di nuove metriche per la valutazione della ricerca;

c) sostenere lo sviluppo e il mantenimento di strumenti, repertori e risorse necessarie al progresso e alla sostenibilità dell'Accesso Aperto (*Software Open Source* e gratuiti, strumenti per il *text-mining* e il *data-mining*, repertori di Archivi e Riviste, Repertori di *Policy* [V. *Report of the European Commission - National Open Access and Preservation Policies in Europe*], di istituzioni e finanziatori della ricerca, fornitori di Licenze Aperte, servizi di "*Preservation and Curation*" [Ball, 2010] di contenuti digitali, servizi per di *cross-linking* e URL persistenti, motori di ricerca).

4. L'importanza di richiamare l'attenzione di Atenei ed Enti finanziatori di ricerca per:

a) trovare modalità di sostegno/patrocinio per le Riviste ad Accesso Aperto che non richiedano il pagamento per la pubblicazione;

b) sostenere la creazione di Standard aperti per i metadati a supporto descrittivo dei contenuti di ricerca che editori di Riviste e gestori di Archivi Aperti possano utilizzare per rendere i contenuti ad Accesso Aperto più facilmente ricercabili e utili.

5. L'importanza della diffusione della conoscenza sulla condotta dell'editoria ad Accesso Aperto (informandoli sulle Licenze, sui processi editoriali, sulla gestione delle quote di pubblicazione).

CAPITOLO 1

Movimento *Open Access*: origini, strategie, politiche e iniziative

La nostra missione di disseminazione della *conoscenza*
è incompleta se l'informazione non è resa *largamente*
e prontamente *disponibile* alla *società*

[V.sito: “Dichiarazione di Berlino sull'accesso
aperto alla letteratura scientifica”, 2003]

1.0. Introduzione

In questo Capitolo sono presentati i presupposti e alcune importanti Dichiarazioni del movimento *Open Access*, il quale si basa su un insieme di azioni internazionali, nate nel contesto accademico-scientifico e riferite

all'ambito digitale, con l'obiettivo di promuovere la massima diffusione e condivisione dell'informazione e quindi la crescita complessiva della conoscenza, incoraggiando scienziati, ricercatori e studiosi a *disseminare* i propri *lavori di ricerca*, *rendendoli liberamente accessibili in rete* [V. sito: “Linee Guida per gli Archivi Istituzionali”, p.5].

In particolare, il movimento *Open Access* si sviluppa su tre fronti:

1) *politico* - con una serie di strategie volte a promuovere le creazione e l'adesione delle istituzioni alle politiche *Open Access*;

2) *tecnologico* - con l' *Open Archives Initiative* (OAI) a supporto delle infrastrutture di Archivi di ricerca (istituzionali e disciplinari) *Open Access* definiti la “via verde” (*Green Road*) dell' Accesso Aperto;

3) *editoriale* – attraverso le Riviste ad Accesso Aperto, “il cui scopo è riguadagnare possesso della comunicazione scientifica”[V. sito “Wiki sull'*Open Access* in Italia”] da parte dei ricercatori e le istituzioni che finanziano la loro ricerca.

Il Capitolo introduce il lettore alla conoscenza dello stato dell'arte delle Politiche *Open Access* che si stanno attivamente promuovendo in diverse parti del mondo. Saranno inoltre messe in luce le varie iniziative internazionali a supporto dell'*Open Access*, incluso l'Italia.

1.1. Presupposti e importanti dichiarazioni sull'*Open Access*

Le attività a sostegno della disseminazione aperta delle informazioni scientifiche risalgono già agli anni Sessanta del ventesimo secolo, quando negli Stati Uniti in ambito biomedico e delle scienze matematiche e fisiche furono

sperimentati i primi sistemi di scambio di fotocopie di *pre-print*² [...] Successivamente si concretizza la disseminazione dei contributi scientifici in versione elettronica utilizzando le potenzialità del web [...] Il *nuovo modello di comunicazione scientifica* nasce negli *anni Novanta* negli Stati Uniti e in altre aree del mondo e oggi si va affermando quale paradigma internazionale del principio per cui lo sviluppo delle conoscenze dipende dalla libertà di coltivarle e farle circolare [Guerrini, 2010, p. 103; Harnad, 1991].

Dagli anni 90 del secolo scorso, i ricercatori hanno iniziato a depositare i loro lavori in Depositi disciplinari ad Accesso Libero come:

1) *Los Alamos* (1991, poi noto come ArXiv [V.sito]), il famoso server di pre-print per i fisici, che ha il merito di aver rivoluzionato per primo la comunicazione dell'informazione scientifica digitale;

² *Pre-print* sono contributi che non sono stati ancora pubblicati da un editore (come opera autonoma o come articolo in una rivista o in un volume miscellaneo) è dunque che non sono stati sottoposti alla revisione dei pari (*peer review*).

2) *Research Papers in Economics* (RePEc, 1997, [V.sito]), la più grande collezione decentrata di documenti ad Accesso Libero per l'economia, trattandosi di una rete che collega diversi Archivi. Nel 1999 Steven Harnad (creatore di *CogPrints Archive*: l'Archivio digitale di Scienze Cognitive) attraverso il suo scritto «*The Optimal and Inevitable for Researchers*» [V.sito]), presenta un possibile mondo della comunicazione scientifica, più equo di quello attuale, con cui ci permette di delineare quella che potremmo considerare la "situazione ottimale" della comunicazione del materiale di ricerca, sulla quale nessun ricercatore non potrebbe obiettare alcunché. Secondo Harnad, le motivazioni per le quali gli autori pubblicano sono dettate dalle necessità di:

- a) comunicare con i propri colleghi;
- b) avanzare nella carriera;
- c) ottenere finanziamenti per ricerche future;
- d) far conoscere le loro ricerche al grande numero di utenti che possono fruirne e citarne, tramite cui ottenere un ampio riconoscimento e prestigio.

È evidente che il soddisfacimento del primo e dell'ultimo requisito, appoggiandosi solo sulla disseminazione della conoscenza scientifica tramite i canali dell'editoria commerciale, può avvenire solo in parte. Il movimento *Open Access* (OA) amplifica i canali della comunicazione dell'informazione scientifica, accentuando la:

1. necessità di riorganizzare il tradizionale circuito della comunicazione dell'informazione scientifica, promuovendo l'applicazione dell'*Accesso Aperto* ai contenuti finanziati con i fondi pubblici;

2. esigenza di disseminare il materiale di ricerca attraverso le Riviste e gli Archivi ad Accesso Libero, sfruttando al meglio le *potenzialità della rete* e favorire un più ampio accesso ai risultati della ricerca, con la prospettiva di "rimuovere ogni barriera economica, legale o tecnica contro l'accesso all'informazione scientifica, ciò al fine di garantire il progresso scientifico e tecnologico a beneficio di tutta la collettività" [De Robio, 2007, p.145].

La nascita del movimento OA, quindi,

non può essere esclusivamente ricondotta alla reazione da parte del mondo accademico al rapido incremento dei prezzi dei periodici e la sua finalità principale non è la riforma dell'editoria scientifica e, in particolare, del *journal publishing* [Guerrini, 2010, p.19; V. sito "*QScience.com*"]

bensì anche la promozione di *modelli alternativi di circolazione della conoscenza scientifica* tramite i luoghi ben strutturati (Riviste e Archivi ad Accesso Aperto) accessibili tramite l'Internet.

Nel 1998 la *Association of Research Libraries* (ARL) negli Stati Uniti ha dato l'origine all'organizzazione *SPARC* (*Scholarly Publishing Academic Research Coalition* [V. sito]), un'alleanza supportata da 800 biblioteche accademiche e di ricerca in tutto il mondo, creata con lo scopo di trovare e promuovere le innovative soluzioni per la disseminazione della conoscenza scientifica insieme con le opportunità offerte dalle nuove tecnologie di rete, sostenendo la strada alternativa della propagazione della conoscenza scientifica all'accesso libero.

Nel 2002, a seguito del successo di *SPARC USA*, nasce l'associazione europea *SPARC Europe* [V.Sito] con il compito di diffondere le nuove strategie di editoria sostenibile per provvedere l'accesso libero all'informazione scientifica in Europa, sede e dominio quasi assoluto dell'editoria scientifica commerciale.

La formalizzazione e la dichiarazione ufficiale del movimento *Open Access* sono avvenute in sede dei vari convegni a livello internazionale:

1. a Budapest (Ungheria, 2002), riconosciuto come il primo raduno storico di fondazione del movimento *Open Access* con il successivo rilascio del Manifesto "*Budapest Open Access Initiative*" [V. sito "BOAI"].

In occasione del 10 ° anniversario di questa iniziativa, nel 2012 la BOAI è stata

riaffermata e completata con una serie di raccomandazioni concrete da completare entro i prossimi dieci anni. Secondo BOAI-2012 [V. sito "BOAI-2012"], nel prossimo futuro l'OA diventerà il metodo predefinito per la disseminazione della ricerca "peer-reviewed" in ogni campo scientifico e in ogni paese.

2. A Berlino (Germania, 2003), con l'edizione della "Berlin Declaration on Open Access to Knowledge in the Science and Humanities" [V. sito]. Essa si costruisce e si basa sulla definizione della conferenza di Budapest, promuovendo l'Open Access e quindi Internet come strumento funzionale alla comunicazione dell'informazione scientifica, sostenendo la transizione verso il paradigma dell'Accesso Aperto elettronico alla letteratura scientifica.
3. A Maryland (Stati Uniti, 2003), con la pubblicazione della dichiarazione "Bethesda Statement sull'Open Access Publishing" [V. sito].
4. A Messina (2004), dove è stato svolto il convegno "Gli Atenei Italiani per l'Open Access: verso l'accesso aperto alla letteratura di ricerca". In questa occasione le 31 Università italiane³ hanno aderito alla cosiddetta "Dichiarazione di Messina" [V.Sito]:
 - a) aderendo alle iniziative internazionali di promozione e sostegno all'Accesso Aperto alla conoscenza e alla comunicazione scientifica, e in particolare alla "Dichiarazione di Berlino";
 - b) favorendo la creazione e l'utilizzo di Archivi Istituzionali aperti Open Access in ogni Ateneo;
 - c) favorendo la pubblicazione di articoli e risultati di ricerca sulle riviste ad Accesso Aperto.
5. A Brisbane (Australia, 2008), con la formulazione di "Brisbane Declaration on Open Access [V.Sito];
6. A Granada (Spagna, 2010), con il rilascio della "Dichiarazione di Alhambra sull'Accesso Aperto".

Le Raccomandazioni BOAI hanno dichiarato le due strategie complementari [Guédon 2004, 2008; Cullen, Chawner, 2011] dell'Open Access:

I. BOAI-1: (*Open Archive – OA, "Green Road", «via verde»*) il *self-archiving* o l'auto-archiviazione⁴ [V. De Robbio, 2003], il deposito da parte dei ricercatori di una copia elettronica dei loro lavori (*pre- e/o post-print*) in *Repositories/Archivi* (Disciplinari o Istituzionali) basati sulla piattaforma tecnologica *Open Archives* a cui sono affidati i compiti della diffusione e dell'interoperabilità dell'informazione scientifica in rete.

II. BOAI-2: (editoria *Open Access – OA, "Gold Road", «via d'oro»*) la pubblicazione di contributi scientifici nelle riviste ad Accesso Aperto (con ISSN e *peer reviewed*), a cui è riservato il ruolo di certificazione e validazione (*QC/C quality control and certification*, a seguito dell'attività di *refereeing*) dei contenuti scientifici.

Per adottare la logica dell'Open Access non è necessario modificare i propri stili di comunicazione dell'informazione scientifica; basta inviare i propri articoli a una rivista ad Accesso Aperto, oppure depositare nell'Archivio Aperto del proprio Ateneo copia dei lavori pubblicati su Riviste Open Access o Riviste tradizionali, secondo quanto previsto dai maggiori editori per quanto riguarda la cessione del *Copyright*.

Quanto alle riviste ad Accesso Aperto, esse:

³ Bologna, Brescia, Calabria, Firenze, Foggia, Genova, Insubria, Lecce, Messina, Milano, Milano Bicocca, Milano Politecnico, Milano San Raffaele, Modena, Molise, Napoli Federico II, Napoli L'Orientale, Napoli Parthenope, Padova, Palermo, Parma, Piemonte Orientale, Roma LUMSA, Roma Tor Vergata, Roma Tre, Siena, Torino, Trieste, Trieste SISSA, Tuscia, Venezia IUAV.

⁴ Per *self-archiving* (lett.: "auto-archiviazione") si intende la pratica con la quale ogni autore pubblica autonomamente il proprio materiale su un Archivio Aperto, attraverso una semplice procedura informatica guidata, Glossario dell'Archivio Istituzionale ad Accesso Aperto dell'Università degli Studi di Torino,

<<http://library.unito.it/doc/Glossario.pdf>>

a) fanno parte di *editoria elettronica sostenibile* (che presume altre fonti di finanziamento per il referaggio o la pubblicazione);

b) non prevedono abbonamento o tasse per l'Accesso online; assicurano una visibilità molto più ampia rispetto a quella che può offrire una pubblicazione su un periodico tradizionale accessibile solo agli abbonati.

Una pubblicazione *Open Access* deve soddisfare due requisiti:

1. garantire a tutti gli utilizzatori il diritto d'accesso libero (accesso gratis e Libre' Open Access [V. sito]) ad essa, autorizzandone la riproduzione, l'utilizzazione, la distribuzione e la trasmissione, purché ne sia riconosciuta la paternità intellettuale.

L'*Open Access* viene presentato con due gradi di libertà diversi: "Gratis OA" è l'accesso online senza costi per l'utente, mentre il "Libre OA" è il "Gratis OA" più alcuni addizionali diritti di sfruttamento [V. sito "*Gratis and Libre Open Access*"]. Spesso, questi diritti vengono garantiti all'utente utilizzando le licenze *Creative Commons* (*Creative Commons-Attribuzione*, *Creative Commons-Attribuzione-Non commerciale* e altri);

2. far sì che il contributo sia depositato in un Archivio digitale (Istituzionale e/o Disciplinare) basato sul protocollo *Open Archive Initiative Protocol for metadata Harvesting* (OAI-PMH) e standard internazionali.

Le due strategie (BOAI-1 e BOAI-2) sono state ritenute complementari in considerazione del fatto che esse, non essendo in competizione, possono migliorare l'efficacia dell'Accesso Aperto alla conoscenza (sia per gli "studiosi-autori" che "studiosi-lettori") [Guerrini, 2010, p. 19] grazie ad un'appropriata combinazione tra loro [Guédon, 2008]:

Sostanzialmente si attua un doppio binario: da una parte il deposito immediato entro un Archivio di *e-print* pubblicamente accessibile (istituzionale o disciplinare), dall'altra la sottomissione alla rivista secondo il processo tradizionale, che costituiscono quindi un fattore di riequilibrio a favore degli autori scientifici e degli utenti, ripristinando un ambiente favorevole alla circolazione delle idee e alla tutela della conoscenza come bene comune [Guerrini, 2010, p.108].

L'eprint (o e-print) assume due significati: da una parte è la versione digitale di un contributo scientifico (per esempio, un articolo di giornale, una tesi di diploma, una relazione a un convegno, un capitolo di un libro), accessibile online, dall'altra, identifica l'archivio aperto tout court [ivi, p.106].

La strategia "*Gold Road*" copre non soltanto articoli di Riviste, ma anche capitoli di libri e monografie, ed in generale ogni prodotto della ricerca [V. sito "*Open Access Overview*"]. Da punto di vista dei modelli, l'*Open Access* può essere espresso anche tramite:

- 1) il *modello editoriale* (Riviste, Monografie), che focalizza il modo alternativo all'editoria tradizionale con cui si realizza la gestione dei diritti di proprietà intellettuale e la funzione di filtro di qualità operata dalla *peer review*.

Un recente rapporto su un progetto olandese "OAPEN-NL: exploring *Open Access* monograph publishing in the Netherlands" ha dimostrato che le monografie OA non solo non producono effetti negativi sulle vendite del libro ma ne incrementano l'uso e la reperibilità online[Ferwerda, Snijder, Adema, 2013]. I risultati illustrati dal medesimo Rapporto forniscono importanti informazioni agli autori, agli editori e ai bibliotecari in merito alle potenzialità delle monografie OA;

- 2) il *modello della Biblioteca Digitale* (incluso Archivi digitali di ricerca), come sistema di selezione, deposito, organizzazione, recupero, preservazione ed uso di contenuti testuali ed informativi (metadati);
- 3) il *modello della gestione di conoscenza* (*Knowledge Management*) che, oltre alle funzioni tradizionali delle Biblioteche, ha al centro nuove funzioni collaborative tra comunità d'apprendimento coinvolte, anche attraverso l'integrazione e l'interoperabilità di network eterogenei, incluso quelle progettate per le comunità di apprendimento [V. sito "*Open Access* e condivisione delle conoscenze"].

Ad ogni modello la filosofia OA applica i cambiamenti in termini:

- a) *economici*, proponendo i nuovi modelli economici di editoria digitale [V.siti: “Modelli economici di finanziamento dei lavori *Open Access*”];
- b) *legislativi*, promuovendo l’applicazione delle Licenze di editoria alternative [V. siti: “Licenze”];
- c) *tecnologici*, sfruttando le potenzialità offerte dal Web e degli Standard di interoperabilità tecnica e semantica) che cambiano anche profondamente le relazioni tra persone che comunicano e riusano la conoscenza in rete.

La dimensione comunicativa e quella collaborativa sono sicuramente gli aspetti più interessanti del fenomeno Internet, quelli su cui conviene soffermarsi più attentamente: in queste dimensioni la rete in quanto tecnologia presuppone in realtà una rete di persone [V. sito: “Semantica e Terminologia nei portali”].

Internet si è posto anche come forza interna dell’OA per la valutazione della ricerca, attraverso l’applicazione di metriche di nuova generazione come il *web impact factor*, la *citation analysis* e l’*usage factor* [V. sito: Progetto “*An Analysis and evaluation*”]. Pubblicare in rete rende poi possibile avvalersi di statistiche di *download* e di *trackback* (statistiche sugli accessi al server, ma anche sul singolo autore e su ogni suo singolo articolo) [V. siti: Database citazionali].

I vantaggi dell’applicazione delle strategie *Open Access* nel campo della ricerca sono i seguenti:

1. aumento dell’impatto e della diffusione dei risultati della ricerca;
2. maggiore consapevolezza e controllo sui diritti dell’autore e del processo editoriale;
3. possibilità di fruire di una modalità di pubblicazione alternativa rispetto al mercato editoriale tradizionale;
4. disponibilità di nuovi strumenti e indicatori per il monitoraggio e la valutazione della ricerca.

I peggiori nemici dell’OA sono i luoghi comuni causati dalla scarsa conoscenza delle problematiche. Spesso i docenti non si interessano adeguatamente ai problemi legati al diritto d’autore, hanno poca dimestichezza con le procedure di valutazione della ricerca scientifica, sottovalutano le potenzialità della comunicazione online. Questi problemi non si affrontano con l’introduzione di un obbligo nei confronti dei docenti a depositare una copia dei propri lavori nell’IR, tentazione presente in molte realtà e che rischia di avere conseguenze controproducenti [Occorre, dunque] convincere l’interlocutore [...] rispondendo a tutte le obiezioni [Guerini, 2010, p.81]⁵

Attualmente, l’Accesso Aperto è all’origine di molte discussioni tra universitari, bibliotecari, amministratori di università e politici. Esiste una sostanziale dissonanza sul concetto dell’*Open Access*, che può significare sia l’accesso libero/aperto ad un’opera, sia il movimento che porta avanti questa strategia. È in corso un grande dibattito circa la remunerazione economica dell’*Open Access* da parte degli universitari e i modelli di business da adottare.

L’accesso aperto non è senza costi e richiede modelli economici che garantiscano la sua sostenibilità come qualsiasi altra forma di editoria. Tuttavia, i suoi benefici in termini di visibilità, estensione e rapidità della disseminazione, compressione del tasso di duplicazione delle ricerche, potenziamento della ricerca interdisciplinare, trasferimento della conoscenza alle imprese, trasparenza verso la cittadinanza sono immensi. Tant’è che diverse istituzioni finanziatrici, università ed enti di ricerca, Stati e organizzazioni internazionali stanno perseguendo con determinazione Politiche per favorire la definitiva affermazione dell’Accesso Aperto [Caso, 2013, p.10].

1.2. Stato dell’arte delle Politiche *Open Access*

In seguito sono illustrati i punti principali sulla visione delle Politiche *Open Access* nell’Unione Europea. Queste informazioni vengono fornite tramite un’infrastruttura a supporto dei ricercatori

⁵ All’Università di Harvard, ad esempio, è stato creato un ufficio che si occupa di assistere i docenti e di fornire loro informazioni su tutti gli aspetti dell’OA. L’ufficio ha un ruolo decisivo anche nel promuovere e sostenere l’attività di *self-archiving*.

europei *OpenAIRE* [V. sito], particolarmente attraverso il documento “*Open Access Policies and Mandates*”:

1) Ci sono notevoli vantaggi economici, sociali e didattici che possono derivare dall’accesso aperto ai risultati della ricerca senza barriere economiche, legali e tecniche. L’*Open Access* viene integrato entro la promozione della ricerca nazionale e internazionale in una rete, aumenta l’impatto della ricerca stessa, fornendo nuove *partnership* di ricerca, e rimuovendo l’isolamento professionale.

2) Il Capitolo 2.5.2. dell’Agenda digitale per l’Europa – “*Driving ICT innovation by exploiting the single market*” - si riferisce alle attività gestite in modo efficace per il trasferimento delle conoscenze, affermando che la ricerca finanziata con fondi pubblici dovrebbe essere ampiamente divulgata attraverso la pubblicazione di contenuti di ricerca (compreso i dati di ricerca) in modalità *Open Access*.

3) L’iniziativa “*Europe 2020 Flagship Initiative Innovation Union*” ha una simile clausola sull’Accesso Aperto: la Commissione Europea intende promuovere l’Accesso Aperto ai risultati della ricerca finanziata con fondi pubblici, particolarmente finanziati dai programmi quadro di ricerca dell’UE.

4) Diversi governi hanno cominciato ad interessarsi per l’*Open Access* con l’obiettivo di garantire che la ricerca che essi finanziano possa raggiungere il più vasto pubblico possibile. Ad esempio, il mandato dell’*Open Access* ai risultati della ricerca della *U.S. National Institutes of Health (NIH)* è passato tramite il Congresso degli Stati Uniti ed è stato firmato dal Presidente Obama.

5) Uno dei primi e più importanti esempi delle Politiche *Open Access* nel mondo di finanziatori della ricerca è fornito dalla fondazione senza scopo di lucro britannica “*Wellcome Trust UK*” che tramite la “*Position statement in support of Open and unrestricted Access to published research*” [V. sito], impone ai ricercatori che ricevono finanziamenti di pubblicare i risultati in ambiente *Open Access*, per garantire la loro più ampia disseminazione. “*Wellcome Trust UK*” ha condiviso a livello globale i risultati scientifici -in modalità *Open Access* - relativi a straordinarie scoperte scientifiche per il miglioramento nel campo della salute.

Le più importanti *Directory* che promuovono la conoscenza sulle politiche istituzionali *Open Access* sono:

1. MELIBEA [V. sito], che elenca 64 finanziatori di ricerca, le cui Politiche sull’Accesso Aperto richiedono ai ricercatori di pubblicare i loro lavori in modalità *Open Access*;
2. SHERPA JULIET [V. sito], che espone una lista di 56 finanziatori di ricerca che regolano i loro rapporti con i ricercatori tramite i mandati e le Linee Guida, su come e dove e dove depositare i lavori di ricerca dopo aver ricevuto un finanziamento pubblico;
3. ROARMAP (*Registry of Open Access Repository Material Archiving Policies*) [V. sito], che enumera 47 organizzazioni - finanziatori di ricerca (in Australia, Austria, Belgio, Canada, Francia, Ungheria, Irlanda, Italia, Norvegia, Spagna, Svezia, Svizzera, Regno Unito e Stati Uniti) che presentano in dettaglio i testi delle Politiche di finanziamento in connessione all’Accesso Aperto.

Alcuni importanti suggerimenti su come sviluppare una Politica *Open Access* per un Ente finanziatore della ricerca si possono trovare entro i seguenti documenti digitali di pubblico dominio:

1. Linee Guida *MedOANet for implementing Open Access Policies* [V. sito];
2. Raccomandazioni “*Peter Suber’s Open Access Policy options*” [V. sito];
3. UNESCO “*Policy Guidelines for the development and promotion of Open Access*” [V. sito];
4. Linee Guida “*Good practices for university Open-Access Policies*” fornita dal progetto *Harvard Open Access Project (HOAP)* [V. sito];

5. *OECD Principles and Guidelines for Access to Research Data from Public Funding* [V. sito];
6. *Report of the European Commission - Public Consultation on Open Research Data* [V. sito];
7. *EUROHORCs' Recommendations on Open Access (OA)* [V. sito];
8. *COMMUNIA Policy recommendation* [V. sito];
9. *World Bank's Open Access Policy* [V. sito];
10. *European Research Council Open Access Policy* [V. sito];
11. *Department for International Development "DFID Open and Enhanced Access Policy"* [V. sito].

Dalle indicazioni in materia fornite precedentemente, si può evincere che l'*Open Access* ha goduto di una grande quantità di accettazione e di crescita, con un numero particolarmente elevato di iniziative per la creazione e l'adozione delle Politiche *Open Access*, incluso anche le famose dichiarazioni a supporto delle Politiche *Open Access*:

- ✓ dell'*International Federation of Library Associations (IFLA) "Statement on Open Access to Scholarly Literature and Research Documentation"* [V. sito];
- ✓ dell'*European University Association "Statement from the EUA Working Group on Open Access"* [V. sito].

Con tale quadro di riferimento diventa evidente che l'Accesso Aperto alla comunicazione dell'informazione scientifica è ormai entrato nel circuito della condivisione della conoscenza a livello globale.

Riagganciandoci alla questione della formulazione delle attività ovvero della Politica (*Policy*) *Open Access* negli Archivi Istituzionali, sarebbe utile consultare il documento "*The optimal institutional Open Access Policy*", sul sito dedicato dell'Università di Liège. Secondo il medesimo documento una Politica istituzionale *Open Access* potrebbe essere formulata attraverso i seguenti punti:

[Nome dell'Istituzione] richiede che autori degli articoli - che descrivono e pubblicano le ricerche finanziate con fondi pubblici - massimizzano l'accessibilità, l'utilizzo e le applicazioni delle loro scoperte scientifiche. A tal fine:

[Nome dell'Istituzione]

- 1) richiede che le copie elettroniche di tutti i materiali di ricerca - supportati in tutto o in parte dai finanziamenti pubblici e accettati per la pubblicazione in una rivista *peer-reviewed* - vengono depositati nell'Archivio digitale Istituzionale, immediatamente dopo l'accettazione per la loro pubblicazione;
- 2) richiede che assieme al deposito del materiale di ricerca nell'Archivio Istituzionale siano creati i metadati (titolo, autore/i, affiliazione istituzionale, nome della rivista che ha pubblicato l'articolo), tramite cui il medesimo materiale può essere trovato ed eventualmente consultato in linea;
- 3) richiede che la consultazione del *full-text* del materiale depositato sia resa disponibile entro e non oltre 6 mesi dopo la pubblicazione del materiale di ricerca in una rivista scientifica;
- 4) incoraggia gli autori (creatori delle opere e quindi titolari dei diritti esclusivi morali e di sfruttamento) a riservare e/o non cedere completamente all'editore la proprietà del Diritto d'autore (*Copyright*) [Ricolfi, 2012] dei propri materiali pubblicati.

La Politica deve essere accompagnata da una spiegazione dei vantaggi dell'Accesso Aperto al materiale di ricerca, sia per i ricercatori che per l'Istituzione, come segue:

Quali sono i vantaggi dell'*Open Access* per i ricercatori?

a) *i ricercatori nel ruolo di autori*, che pubblicano i loro lavori di ricerca in modalità *Open Access*, traggono i seguenti benefici: i medesimi lavori hanno una diffusione molto più ampia e possono essere letti senza restrizioni da chiunque abbia accesso a Internet. Questo aumenta l'impatto della loro ricerca: gli articoli ad Accesso Aperto sono citati in 25-25% di più rispetto agli articoli ad accesso chiuso o ristretto delle Riviste in rete;

b) *i ricercatori nel ruolo di lettori* saranno in grado di accedere e utilizzare il testo completo (*full-text*) di un gran numero delle ricerche pubblicate nella loro area scientifica.

Quali sono i vantaggi dell'Open Access per l'Istituzione [nome]?

1) In primo luogo, la ricerca del *[Nome dell'Istituzione]* sarà più accessibile ai ricercatori a livello mondiale, e quindi più conosciuta e più ampiamente utilizzata e citata.

Il prestigio di ricercatori di alto profilo del *[Nome dell'Istituzione]* verrà incrementato. Anche i ricercatori meno noti guadagneranno in termini di esposizione e di impatto dei loro risultati di ricerca in modalità *Open Access* tramite l'Archivio Istituzionale del *[Nome dell'Istituzione]*.

2) In secondo luogo, tutta la ricerca del *[Nome dell'Istituzione]* sarà aperta a tutto il pubblico con accesso all'Internet.

Questo presenterà un vantaggio sia economico che culturale. Dati di terzi, accesso, utilizzo e citazione della ricerca esposta tramite l'Archivio Istituzionale saranno sempre disponibili e analizzabili, fornendo le basi per la formulazione delle Politiche e strategie di ricercatori, istituzionali e quelle nazionali.

Che cosa dovrei depositare quando ho un lavoro pronto per la pubblicazione?

Si deve depositare la versione finale del materiale di ricerca, così come è stato accettato e pubblicato nella rivista scientifica.

Quando devono essere depositati i materiali di ricerca nell'Archivio?

Una versione elettronica della versione finale della pubblicazione dell'autore derivante dalla ricerca - sostenuta, in tutto o in parte, da finanziamenti pubblici - deve essere depositata immediatamente dopo l'accettazione per la pubblicazione.

Come posso sapere se la Rivista in cui ho pubblicato il mio lavoro di ricerca ha una politica compatibile con il deposito del medesimo lavoro nell'Archivio Istituzionale?

Si consiglia di consultare le Politiche editoriali di diverse riviste scientifiche sul sito del SHERPA/RoMEO (*Rights on METadata for Open archiving*):

<<http://www.sherpa.ac.uk/romeo/>>. Nel sito viene evidenziato che circa due terzi delle Riviste permettono la pratica di deposito dei lavori pubblicati in esse negli Archivi Istituzionali.

Se il mio lavoro di ricerca è stato già reso ad Accesso Aperto tramite la pubblicazione in una Rivista scientifica Open Access, lo devo comunque depositare nell'Archivio Istituzionale?

Sì, il *[Nome dell'Istituzione]* richiede che i lavori pubblicati (sia nelle Riviste tradizionali che in quelle *Open Access*) siano depositati nell'Archivio Istituzionale dell'Ateneo, per registrarli come un completamento della ricerca istituzionale.

Le applicazioni delle Politiche *Open Access* negli Archivi Istituzionali si possono differenziare nel campo oggettivo, ad esempio,

Mentre Liège applica in modo generalizzato la sua *Policy* a *tutti i generi letterari*, Harvard sceglie di limitare l'applicazione della politica ai soli *articoli scientifici*. Dalla documentazione a disposizione risulta altresì che sia solo *Harvard a prevedere un doppio livello di Policy*: istituzionale (con indicazioni valide per l'intero Ateneo) e sub-istituzionale (con singole *Policy* riguardanti le strutture decentrate) [Caso, 2013, p.30].

In Italia, a seguito di riforma del sistema universitario (c.d. legge Gelmini), la CRUI (Conferenza di Rettori delle Università Italiane)⁶ ha steso una clausola-modello che potrebbe rappresentare la base della Politica *Open Access*, la quale si propone di inserire nei nuovi Statuti di Atenei [V. sito: Clausola-modello per la Politica *Open Access* degli Atenei]:

“1. L’Università di [Nome] fa propri i principi dell’accesso pieno e aperto alla letteratura scientifica e promuove la libera disseminazione in rete dei risultati delle ricerche prodotte in ateneo, per assicurarne la più ampia diffusione possibile.
2. L’Università, con apposito regolamento [da emanare entro centottanta giorni dall’entrata in vigore del presente Statuto], pone la disciplina finalizzata a dare attuazione ai principi dell’accesso pieno e aperto ai dati e ai prodotti della ricerca scientifica, incentivandone il deposito nell’archivio istituzionale e la comunicazione al pubblico, nel rispetto delle leggi concernenti la proprietà intellettuale, la riservatezza e la protezione dei dati personali, nonché la tutela, l’accesso e la valorizzazione del patrimonio culturale”.

La medesima “Clausola-modello” è stata accolta da numerose Università: “alla fine del 2012 - 35 Atenei risultavano aver inserito una clausola sull’Accesso Aperto nei nuovi statuti approvati dal MIUR” [V. sito *Open Access* in Italia 2012].

Una *Policy* Istituzionale universitaria sull’Accesso Aperto deve, inoltre, puntare innanzitutto a investire nel sistema [...] nelle strutture istituzionali e amministrative di riferimento [...] Molte Università italiane sono dotate di sistemi bibliotecari (con relative commissioni) nonché di uffici di ricerca e trasferimento tecnologico (con relative commissioni). Servono altresì strutture istituzionali e amministrative che coordinino le politiche in materia di Accesso Aperto. [Caso, 2013, p.55].

1.3. Strumenti, iniziative e attività a supporto del movimento *Open Access*

Nell’ultimo decennio il movimento *Open Access* si è dimostrato abbastanza forte e attivo nel campo della ricerca scientifica, contribuendo all’apertura di una moltitudine di nuovi canali digitali secondo le strategie *Open Access* (Riviste ad Accesso Aperto [Sutton, 2011] e Archivi Istituzionali e Disciplinari) e in conformità alle Politiche *Open Access*, che vengono adottate e condivise da diverse istituzioni che supportano la produzione scientifica.

Tra gli Archivi aperti *Open Access* di rinomanza a livello internazionale vi sono i già citati Archivi Disciplinari per la disseminazione dei lavori:

- a) di fisici e matematici - ArXiv.org;
- b) di economisti - PubMed;
- c) nel campo delle scienze dell’informazione - E-LIS [De Robbio, Katzmayer, 2009].

Tra le più importanti *Directory* che indicizzano Riviste e Archivi *Open Access* e loro contenuti ci sono:

⁶ È organismo di rappresentanza di 80 università e istituti universitari italiani (<<http://www.cruai.it/>>) che svolge la propria attività anche attraverso il supporto della Fondazione CRUI per le Università italiane (<<http://www.fondazionecruai.it/>>).

Tra i propri compiti istituzionali CRUI ha, tra gli altri, anche quello di svolgere attività di coordinamento, indirizzo, tutela e promozione delle Università Italiane, in un’ottica di sistema. Dal 1999 ha attivato al proprio interno una Commissione dei delegati rettorali per le biblioteche che si è sempre proposta come una sede istituzionale di confronto e sviluppo del sistema delle biblioteche universitarie, per favorirne la crescita e rappresentarne le esigenze.

La Fondazione CRUI per le Università italiane, nata nel 2001 come strumento operativo e progettuale della CRUI, è sempre più un’agenzia al servizio dell’innovazione universitaria. La forte integrazione tra CRUI e Fondazione garantisce un processo continuo di studio, progettazione e verifica dei risultati, per un supporto costante alle 80 Università associate. Da anni impegnata nella promozione e nell’adozione di modelli migliorativi per il sistema, la Fondazione ha sviluppato una vasta esperienza nel coordinamento e nella valorizzazione delle attività degli Atenei, riuscendo a costruire una rete accademica sempre più solida e organizzata.

- 1) DOAJ (*Directory of Open Access Journals*) [V. sito];
- 2) OpenDOAR (*Directory of Open Access Repositories*) [V. sito];
- 3) ROAR (*Registry of Open Access Repositories*) [V. sito];
- 4) *Open Access CORE (CONnecting REpositories)* [V. sito];
- 5) NDLTD (*Networked Digital Library of Theses and Dissertations*) [V. sito];
- 6) DOAB (*Directory of Open Access Books*) [V. sito], la Directory dedicata alle monografie *peer reviewed* rese disponibili tramite licenze *Open Access*;
- 7) Portale NARCIS (*National Academic Research and Collaborations Information System*) [V. sito];
- 8) Portale DART-Europe Portal [V. sito] di ricerca delle tesi di dottorato *full-text* provenienti da 300 Università di 19 paesi europei;
- 9) Portale di ricerca DRIVER (*Digital Repository Infrastructure Vision for European Research*) [V. sito];
- 10) motore di ricerca BASE (*Bielefeld Academic Search Engine*) [V. sito];
- 11) motore di ricerca OAIster [V. sito];
- 12) motore di ricerca *Open Access Thesis and Dissertations* [V. sito] di circa 2 milioni di tesi e dissertazioni da oltre di 600 università di tutto il mondo),

Tra le più notevoli comunità online che promuovono la conoscenza sul movimento *Open Access* si segnalano i seguenti Portali:

1. UNESCO's *Global Open Access Portal (GOAP)* [V.sito];
2. SPARC (*Scholarly Publishing and Academic Resources Coalition*) e SPARC Europe [V.siti];
3. *OpenAIRE (Open Access infrastructure for Research in Europe)* [V.sito];
4. COAR (*Confederation of Open Access Repositories*) [V.sito];
5. EIFL (*Knowledge without boundaries*) [V.sito];
6. *EnablingOpenScholarship (EOS)* [V.sito];
7. *Public Library of Science (PLoS)* [V.sito];
8. PLEADI (la piattaforma nazionale per l'accesso centralizzato alla letteratura scientifica depositata negli Archivi Aperti Italiani, istituiti e gestiti da Università ed Enti di ricerca) [V.sito].

Nel 2009 vengono lanciate cinque nuove collane di monografie culturali umanistiche *Open Access* dai titoli: “*New Metaphysics*”, “*Critical Climate Change*”, “*Global Conversations*”, “*Unidentified Theoretical Objects*”, e “*Liquid Books*” [V.sito AIDALampi].

Nel 2011 dalla Rice University sono stati lanciati i progetti “*OA books*” per la letteratura dell’area umanistica [V.sito; Altman Micah, 2013].

Alla fine del 2011 vengono registrati

over 7,000 peer-reviewed fully Open Access Journals as listed in the DOAJ, still growing by 4 titles per day and over 6,000 of these are in English, as listed by Open J-Gate. Electronic Journals Library keeps track of more than 32,000 free Journals. There are over 2,000 Repositories, linking to more than 30 million items, growing at the rate of 21 thousand items per day, which can be searched through the snazzy new Bielefeld Academic Search Engine search options. PLoS ONE, having become the world's largest Journal last year, outdid themselves by doubling the number of articles published this year. PubMedCentral, arXiv, RePEC, and E-LIS growth was in the 10-15% range for the year. This issue of Dramatic Growth adds a new feature, a first attempt at comparing compliance rates with a few medical funders' open access policies - so far, Wellcome Trust is looking good [V. sito “*The Dramatic Growth of Open Access*”].

La British Academia ha preparato una ricca riflessione sulle sfide affrontate dalle pubblicazioni *Open Access* nelle Scienze Umanistiche e Sociali [Debating Open Access, 2013].

La JISC (*Joint Information Systems Committee*) insieme con l’*Arts and Humanities Research Council (AHRC)* hanno finanziato il Progetto britannico OAPEN-UK (2010-2015) [V. sito] per la pubblicazione delle monografie *Open Access*.

Nel 2012 la Commissione Europea ha pubblicato le “*Raccomandazioni sull’Accesso all’informazione scientifica e sulla sua conservazione*” (417 del 17 luglio 2012), in cui viene indicato che lo sviluppo di un’economia basata sull’innovazione e sull’ampia disseminazione della conoscenza scientifica rappresenta una priorità:

Tra le azioni da avviare nell’ambito della «Agenda digitale», è opportuno dare ampia diffusione alla ricerca finanziata con fondi pubblici mediante la pubblicazione di dati e studi scientifici ad Accesso Aperto [V. sito: “Raccomandazioni sull’Accesso”].

Diverse nuove tendenze e proposte relative al movimento *Open Access* sono affiorate durante:

- a) le annuali (dal 2005) conferenze “*Munin Conference on Scientific Publishing*” [V. sito] così come anche durante
- b) le annuali (dal 2007) settimane “*Open Access Week*” [V. sito] programmate a livello internazionale. Durante la Settimana internazionale OA (Berlino 11) del 2013, ad esempio, è stato rilasciato un utile strumento per la ricerca alternativa in modalità *Open Access* ovvero per una “*immediate online availability of research articles, along with the rights to use these articles fully in the digital environment*” [V. sito *Open Access Button*] di articoli scientifici con accessi negati sul web.

Dal 2007 l’*European Research Council* (ERC) ha richiesto che tutte le pubblicazioni *peer reviewed* prodotte nell’ambito di progetti finanziati dall’ERC [V. sito: “Pubblicazioni dei progetti finanziati dall’ERC”] siano depositate - immediatamente dopo la loro pubblicazione - in un Archivio *Open Access*, e rese liberamente accessibili - immediatamente o al più tardi 6 mesi (il periodo di embargo) dopo la loro pubblicazione.

Per appoggiare in parte tale richiesta, nel 2010 il Progetto OpenAIRE (2010-2013) ha lanciato OpenAIRE *Orphan Repository* [V. sito], un Archivio aperto digitale *Open Access* in cui possono essere depositate le pubblicazioni dei ricercatori le cui Istituzioni non possiedono ancora un proprio Archivio *Open Access*.

Lo stesso Progetto – per offrire uno spazio informativo per ricercatori, enti e finanziatori di ricerca per quanto riguarda le questioni su:

- a) Accesso Aperto,
- b) Diritto d’autore,
- c) Archivi ad Accesso Aperto -

ha creato una rete di servizi online a livello internazionale - *National Open Access Desk* [V. sito].

Entro il progetto *OpenAIREplus* [V. sito] - che ha lavorato in tandem con l’*OpenAIRE* fino a 2013 – sono stati sviluppati i *collegamenti* tra diverse categorie di pubblicazioni di ricerca, dati di ricerca e i progetti di finanziamento.

Nel 2012 (n.417 del 17 luglio) la Commissione Europea pubblica le “*Raccomandazione sull’Accesso all’informazione scientifica e sulla sua conservazione*” [V. sito], chiedendo alle istituzioni impegnate nella ricerca scientifica di elaborare precise politiche per attuare il principio dell’Accesso Aperto.

Nel 2013 viene lanciato il *Repository Zenodo* [V. sito], l’Archivio digitale dei Dati Aperti di ricerca - frutto della collaborazione tra CERN e il citato progetto OpenAIRE. *Zenodo* presenta un database per il deposito e la conservazione dei risultati di ricerca da parte di tutti i ricercatori, singoli o parte di gruppi di ricerca, comunità, o Istituzioni non dotate ancora di un Archivio dei contenuti di ricerca. I contenuti nell’Archivio *Zenodo*

vengono affinati da opportuni metadati esposti sul Web via protocollo standard dell’interoperabilità *Open Archive Initiative – Protocol for Metadata Harvesting* (OAI-PMH) [V. sito], di modo che i metadati stessi siano facilmente individuabili e recuperabili per il libero uso, scambio, riuso e distribuzione da parte di tutti gli studiosi delle comunità scientifiche [...] Attraverso i metadati [...] *Zenodo* comunica con i servizi Web 2.0 [...] colloquia infatti anche con Mendeley, servizio online recentemente acquisito da Elsevier, che utilizza la tecnologia *Cloud* per la creazione di bibliografie a scopo citazionale [De Robbio, 2013, p.1].

Sempre nel 2013 LIBER (*Ligue des Bibliothèques Européennes de Recherche – Association of European Research Libraries*) [V.sito], COAR (*Confederation of Open Access Repositories*) [V.sito] e il Progetto OpenAIRE hanno pubblicato uno *Statement* congiunto sull'*Open Data Pilot* [V.sito] del programma HORIZON 2020 promosso dalla Commissione Europea. Il medesimo Programma accentua l'importanza di disseminazione dei risultati della ricerca attraverso i canali *Open Access*.

Il Progetto europeo MedOAnet (*Mediterranean Open Access Network*) ha rilasciato nel 2013 le "Linee Guida per l'implementazione delle Politiche *Open Access*", a supporto degli enti di ricerca e dei finanziatori di ricerca [V. sito LINEE GUIDA: *MedOANet for implementing Open Access policies*].

1.4. Open Access in Italia

L'Italia partecipa al movimento *Open Access* con forte impegno, anche a livello internazionale in quanto è coinvolta in diversi Progetti sull'*Open Access* [V. sito: *Wiki.Openarchives.it*].

Nel 2006 l'Associazione Italiana Biblioteche (AIB) [V. sito] ha espresso la sua posizione a favore dell'Accesso Aperto alla letteratura scientifica. Nello stesso anno, all'interno della Commissione dei Rettori delle Università Italiane (CRUI) è stato attivato un gruppo di lavoro *Open Access* (Commissione delle Biblioteche), con il mandato di approfondire i temi legati all'Accesso Aperto, in sinergia con le Politiche e gli indirizzi europei. Questo gruppo di lavoro, individuando diverse *Best Practices*, nazionali e internazionali, ha sviluppato varie Linee Guida da proporre alle Università italiane e agli organi gestionali della CRUI per la implementazione pratica dei principi fondamentali dell'*Open Access*. Le comunicazioni del Gruppo – relative alle novità, iniziative in corso e proposte sugli sviluppi futuri della libera disseminazione dei risultati delle ricerche condotte nelle Università e nei centri di ricerca italiani - sono visibili in un luogo virtuale di dibattito - *Blog "240inpratica"* [V. sito].

Nel 2007 la Commissione Biblioteche CRUI ha approvato le "Linee guida per il deposito delle Tesi di Dottorato negli Archivi Aperti" [V. sito], che forniscono agli Atenei suggerimenti su come "regolare" le procedure collegate al deposito (diritto d'autore; previsione di un eventuale embargo; implementazione di una serie di procedure regolamentari e operative).

Nello stesso anno è stato avviato il processo di semplificazione della procedura di deposito legale delle Tesi di Dottorato presso le due Biblioteche Nazionali Centrali (di Firenze e Roma), introducendo la possibilità della loro consegna in formato elettronico e per via telematica, direttamente dagli Archivi Istituzionali, ove possibile⁷.

Nel 2008 dal Gruppo CRUI *Open Access* sono state tradotte le "*Recommendations from the EUA Working Group on Open Access adopted by the EUA Council*" [V. sito], rivolte alle Università, alle Conferenze dei Rettori, e alla EUA (*European University Association*). Tramite questo documento alle Università Italiane viene raccomandato di:

- a) implementare Archivi (*Repository*) Istituzionali di ricerca e/o partecipare a *Repository* condivisi che aderiscono al protocollo OAI-PMH che consente l'interoperabilità tra gli Archivi finalizzata ad un utilizzo sempre maggiore dei contenuti in rete;
- b) definire Politiche istituzionali in base alle quali i ricercatori afferenti siano incentivati a depositare le loro pubblicazioni scientifiche (attraverso la procedura dell'auto-

⁷ Fino al 2006 il Deposito delle Tesi del Dottorato avveniva in formato esclusivamente cartaceo, con un accesso regolato da forti restrizioni. Dal momento che le Tesi nascono in formato elettronico, e vista la propensione delle Università a far confluire tali documenti nei propri Archivi Istituzionali, già nel 2006 era stata prevista la possibilità di depositarli su supporto elettronico – DVD o CD-ROM (Nota MiUR del 28 luglio 2006, n. 1420). Nel 2007 è stato raggiunto un Accordo con il MiUR (Nota MiUR del 20 luglio 2007, n. 1746), in base al quale le Università potranno consegnare tali Tesi di Dottorato in formato elettronico e per via telematica, per motivi di razionalizzazione e di standardizzazione delle procedure, nonché per snellire il lavoro di catalogazione.

archiviazione) nei *Repository* Istituzionali al momento dell'accettazione della loro pubblicazione;

- c) sviluppare le Politiche istituzionali e strategie che favoriscano la reperibilità dei risultati di ricerca da parte di un bacino di utenti più ampio possibile, massimizzando la visibilità, l'accessibilità e l'impatto scientifico dei risultati stessi (nel nominare a titolo esemplificativo la Politica istituzionale *Open Access to Scientific Publications* dell'Istituto Superiore di Sanità) [De Castro, 2008].

Nel 2009 dal medesimo Gruppo CRUI vengono pubblicate:

le "Linee guida per gli Archivi Istituzionali" [V. sito], che riportano le modalità organizzative e tecniche per la creazione e gestione degli Archivi, in funzione delle loro finalità istituzionali.

Ormai un elevato numero delle Università Italiane si è dotato di Archivi Istituzionali rispondenti alle medesime "Linee". Il già menzionato Service Provider nazionale per l'accesso alla letteratura scientifica prodotta in Italia PLEIADI allo stato del 2013 ha indicizzato oltre 605.000 risorse digitali provenienti dagli Archivi e le Riviste ad Accesso Aperto [V. sito: "*Wiki.Openarchives.it: Open Access in Italia*";

- 1) le "Raccomandazioni in materia di valutazione della ricerca" [V. sito], in cui si sottolinea l'importanza di visibilità e di connesso impatto scientifico della ricerca, anche grazie alle nuove metriche della valutazione emergenti;
- 2) le "Linee guida per le Riviste elettroniche ad Accesso Aperto" [V. sito]. Le Riviste elettroniche ad Accesso Aperto sono uno strumento valido, a costi ragionevoli, per favorire la disseminazione dei risultati della ricerca scientifica.

Nel 2012:

- a) vengono pubblicate le "Linee guida per la creazione e la gestione di metadati nei Repository istituzionali" [V. sito], con l'obiettivo di suggerire agli Atenei uno schema di metadati (indici descrittivi) il più possibile interoperabile nell'ambito di gestione degli Archivi Istituzionali italiani;
- b) il gruppo italiano del Progetto europeo MedOAnet (*Mediaterranean Open Access Network*) [V. sito] - un progetto europeo volto al miglioramento di strategie, politiche e strutture per l'accesso aperto in sei Paesi del Mediterraneo: Italia, Grecia, Francia, Spagna, Portogallo e Turchia - ha trasmesso al MIUR una "Proposta di azione normativa sull'Accesso Aperto in Italia", che richiama ad una necessaria azione governativa che definisca la posizione dell'Italia in materia di Accesso Aperto ai prodotti della ricerca finanziata con fondi pubblici, per consentire un migliore allineamento con le strategie europee e internazionali.

Nel 2013:

- a) è stata firmata la Dichiarazione "*Position statement sull'Accesso Aperto ai risultati della ricerca scientifica in Italia*" [V. sito] dai rappresentanti della CRUI e degli enti pubblici di ricerca coinvolti nel progetto MedOAnet, per promuovere l'istituzione di Archivi liberamente consultabili e per incoraggiare i ricercatori a rendere disponibili i loro risultati;
- b) è stato firmato dal MIUR il documento "HORIZON 2020 Italia" (HIT 2020) [V. sito], che innesci una profonda rivisitazione degli strumenti, dell'impostazione e delle modalità attraverso cui verrà redatto il prossimo *Piano Nazionale della Ricerca*, enunciando l'importanza delle strategie per la disseminazione dei risultati della ricerca attraverso le strategie *Open Access* (Archivi e Riviste Open Access);
- c) è stato lanciato il nuovo Portale della ricerca italiana "*ResearchItaly*" [V. sito], in cui *Open Access* e *Open Data* sono stati indicati come strumenti per la comunicazione e la condivisione della ricerca scientifica;
- d) è stata redatta la Dichiarazione dei Ministri del G8 (firmata anche dalla Ministra Carrozza), che appoggia pienamente l'idea di Accesso Aperto ai dati della ricerca (punto

- 3) e di libera circolazione dei risultati delle ricerche (punto 4) [V. sito: “*Government news G8*”];
- e) sono iniziati a profilarsi all'orizzonte delle prime Tesi del 26° ciclo di Dottorato, quelle con obbligo di deposito ad Accesso Aperto (conciliando l'Open Access della Tesi con la possibilità di oscurare parti della tesi stessa, in caso di presenza di dati tutelati da segreto industriale) negli Archivi Istituzionali, contemporaneamente con la riscrittura dei Regolamenti dei singoli corsi di Dottorato (secondo DM del 8 febbraio, art. 14) dai nuovi Dipartimenti, come segue:
- “Il Ministero cura la costituzione ... di una banca dati delle Tesi di Dottorato ... Entro trenta giorni dalla discussione e approvazione della Tesi l'Università o il soggetto promotore depositano copia della Tesi, in formato elettronico, nella Banca Dati Ministeriale. Previa autorizzazione del collegio dei docenti, possono essere rese indisponibili parti della Tesi in relazione all'utilizzo dei dati tutelati da segreto industriale ai sensi della normativa vigente in materia. Resta fermo l'*obbligo di deposito della tesi* presso le biblioteche nazionali di Roma e di Firenze”.
- f) è stato redatto un Decreto-Legge 8 agosto 2013, n. 91 “Disposizioni urgenti per la tutela, la valorizzazione e il rilancio dei beni e delle attività culturali e del turismo” (GU n.186 del 9-8-2013), secondo cui le pubblicazioni che documentano i risultati di ricerche finanziate per una quota pari o superiore al cinquanta per cento con fondi pubblici devono essere depositate, non oltre sei mesi dalla pubblicazione, in Archivi elettronici Istituzionali o di settore, predisposti in modo tale da garantire l'Accesso Aperto in formato elettronico. L'attuale Governo ha promulgato la legge 122 del 7-10-2013 e ha introdotto nell'ordinamento italiano il principio dell'Accesso Aperto ai risultati della ricerca scientifica finanziata con denaro pubblico [V. sito Legge 7 ottobre 2013, n.112; Caso, Galimberti, 2013];
- g) in occasione della Quinta edizione della Settimana internazionale dell'Open Access (dal 21 al 27 ottobre 2013) [V. sito], in Italia sono state svolte diverse giornate informative sull'evoluzione e sulle iniziative in corso per la promozione dell'Accesso Aperto a livello inter/nazionale, che hanno visto una buona partecipazione di docenti, ricercatori, bibliotecari e utenti comuni. Il 25 ottobre 2013 al CNR di Roma è stato svolto un importante convegno “*E-science: verso un network italiano per l'Open Access*” [V. sito], avente l'obiettivo di condividere i risultati raggiunti nell'ambito di MedOANet (*Mediterranean Open Access Network*) [V. sito]. I risultati dell'evento sono stati promossi anche tramite il nuovo portale MIUR per la ricerca italiana RESEARCH ITALY [V. sito].

1.5. Conclusioni

In questo Capitolo sono stati presentati i presupposti e le importanti Dichiarazioni relative al movimento *Open Access* per la comunicazione dell'informazione scientifica che hanno fornito le ottime basi per:

1. la formalizzazione e l'attuazione di un gran numero di Politiche istituzionali per la divulgazione del materiale di ricerca tramite le strategie *Open Access* da parte di diverse istituzioni ed enti finanziatori di ricerca;
2. la creazione di Riviste e Archivi (Istituzionali e Disciplinari) di ricerca *Open Access* in tutto il mondo;
3. uno sviluppo continuo di diverse ambiziose iniziative a livello nazionale, europeo e internazionale, che riguardano la promozione delle conoscenze sui vantaggi del movimento *Open Access*, la scrittura di Piani, Raccomandazioni, Linee Guida, Norme giuridiche per valorizzare il materiale di ricerca tramite le strategie *Open Access*.

CAPITOLO 2

Archivi aperti *Open Access*: lo stato dell'arte, le questioni di etica e qualità

2.0. Introduzione

Nell'ultimo decennio, in particolare dopo il primo raduno storico di fondazione dell'*Open Access* che ebbe luogo a Budapest, nelle Università e nei centri di ricerca in tutto il mondo sono stati implementati diversi Archivi aperti *Open Access*, a sostegno della proliferazione della “via verde” (“*Green Road*”) dell'Accesso Aperto, come indicato nella Dichiarazione di Budapest BOAI, di cui nel Capitolo precedente.

Il presente Capitolo ha l'obiettivo di mettere in luce alcune importanti tematiche riconducibili all'espressione pratica degli Archivi aperti di ricerca *Open Access*, tra cui:

- h) progetto *Open Archive Initiative* a supporto dell'infrastruttura tecnologica di Archivi aperti;
- i) Archivi Istituzionali del materiale di ricerca, letti alla luce degli argomenti come:
- j) valutazione della ricerca;
- k) etica e proprietà intellettuale dei contenuti di ricerca;
- l) conservazione a lungo termine dei contenuti di ricerca;
- m) *Data (and Process) Management Plan (DMP)* per programmare la gestione di qualità dell'intero ciclo di vita dei contenuti digitali;
- n) *Trusted Digital Repository (TDR)*, un presupposto importante per la certificazione di qualità di piattaforme (Archivi) di gestione dell'informazione digitale.

2.1. Progetto *Open Archive Initiative* a supporto delle piattaforme “Archivi aperti”

L'*Open Archives Initiative* (OAI) [V. sito] è un progetto nato nel 1999 per rendere facilmente fruibili le piattaforme di Archivi che contengono documenti prodotti in ambito accademico.

Il mondo accademico statunitense ha cominciato a utilizzare il termine *Archive* o l'espressione *Open Archives* per indicare quelle collezioni ospitate in server gestiti da Università e da altri istituti, che si arricchiscono dei *risultati delle ricerche in corso* e dei *prodotti di ricerche concluse*, lavori rivolti principalmente ad altri studiosi e liberamente accessibili a tutti tramite il Web [Guerrini, 2010, p.9].

Secondo il «Glossario delle organizzazioni internazionali e dei termini più comunemente utilizzati negli *Open Archives*» [V. sito]: “*Open Archive* è un Archivio di lavori scientifici in formato elettronico organizzati con *metadati descrittivi, liberamente accessibile in rete*”.

L'OAI promuove:

- 1) *Software Open Source* basati sul protocollo di interoperabilità OAI-PMH (*Protocol for Metadata Harvesting*) [V. sito], al fine di consentire la creazione di servizi per:
 - a) rendere utilizzabili tutti i contenuti dell'attività di ricerca degli Atenei attraverso le piattaforme digitali di Archivi aperti;
 - b) abilitare queste piattaforme a interagire tra loro (l'interoperabilità tecnica) e rendere possibile una ricerca su più Archivi simultaneamente;
- 2) lo sviluppo di Standard aperti a supporto della gestione di Archivi;
- 3) la creazione di metadati di qualità regolati da professionisti dell'informazione;
- 4) la creazione di servizi di valore aggiunto per le biblioteche digitali;
- 5) i metodi dell'integrazione delle risorse nel contesto della biblioteca ibrida e lo sviluppo di Portali tematici e istituzionali.

Il modello OAI ha la possibilità di fornire informazione specializzata a vari livelli, catturando i metadati a livello dai diversi Archivi tramite lo schema XML (*eXtensibleMarkupLanguage*) [V. sito]. Tale schema – riconosciuto come uno standard del W3C (Consorzio World Wide Web) per la strutturazione di dati atti a facilitare la condivisione/scambio tra moduli applicativi in

forma sintatticamente e semanticamente non ambigua - specifica quali sono i diritti di utilizzo dei metadati a livello di ogni record (scheda descrittiva della risorsa informativa digitale) e dell'intero Archivio, riferendosi in primo luogo ai metadati descrittivi *Dublin Core* (DC) semplice (*simple*, non qualificato) costituito da 15 elementi per la descrizione di una risorsa [V. Appendice I]. Accanto alla complessità di formati di metadati completamente differenti (es. metadati strutturali METS), i metadati descrittivi DC *simple* vengono scelti come un “*mandatory element set*” o il denominatore minimo comune a supporto delle funzionalità del protocollo OAI-PMH.

Il modello OAI può essere descritto a due livelli:

I. Modello funzionale (Figura 5):

1) *Data Provider* (DP), o *Web Servers Repositories* che gestiscono uno o più Archivi (*Repositories*) di collezioni di oggetti digitali e sono responsabili del loro mantenimento, accesso e della generazione dei metadati che li caratterizzano;

2) *Service Provider* (SP), che possono essere *Harvesters, Aggregators, Caches, Proxies, Gateways*. SP confezionano servizi di ricerca a valore aggiunto, basandosi sull'aggregazione e l'indicizzazione di metadati raccolti dagli Archivi dei DP attraverso le richieste del protocollo OAI-PMH sfruttando le migliori potenzialità del digitale.

Ogni nuova installazione di “Archivi aperti” (DP e SP) - per essere ricercabile, ritrovabile e visibile a livello globale in rete - deve essere opportunamente registrata nel registro internazionale presso il sito OAI per ottenere “OAI:identificatore_archivio”.

Il protocollo dell'interoperabilità, da una parte favorisce la comunicazione, dall'altra limita gli elementi interoperabili a quelli indicati dallo schema DC non qualificato, con il risultato di un certo appiattimento dei risultati della ricerca o di innalzamento della quantità di rumore, con il vantaggio, tuttavia, di permettere ai gestori dei *Data Provider* di condividere i propri dati senza uno sforzo organizzativo gravoso, che in molti casi sarebbe fin troppo penalizzante, limitando così la circolazione delle risorse [Guerrini, 2010, p. 41].

Considerando tale limitazione, sarebbe opportuno un

Service Provider che abbia l'obiettivo di raccogliere metadati da Archivi diversi [riuscisse] a trovare il corretto equilibrio attraverso una mappatura tra differenti stili di indicizzazione per Soggetto o uso di Schemi di Classificazione [De Robbio, 2007, p. 147].

Un buon esempio di un SP è il Portale unificato degli Archivi aperti nazionali di ricerca NARCIS [V. sito], che offre l'Accesso Aperto al maggior numero dell'informazione scientifica aggregata dal Portale prodotta dal mondo accademico e dalle istituzioni scientifiche olandesi (Figura 6).

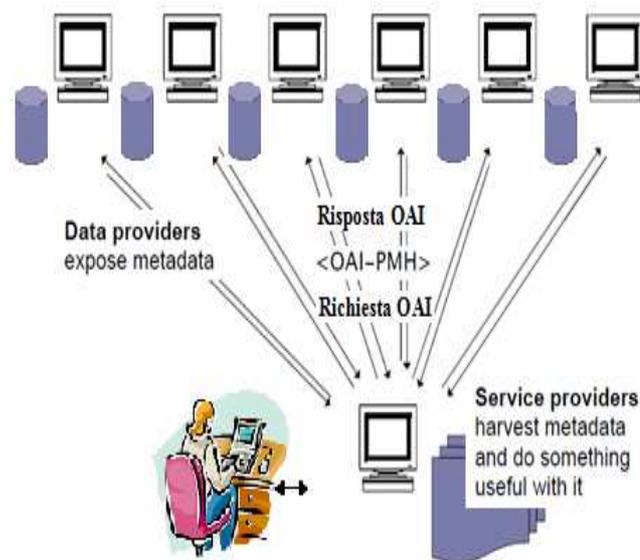


Figura 5: Il modello funzionale *Open Archive Initiative: Data e Service Provider*



The gateway to scholarly information in the Netherlands
 > Submit Content to NARCIS

Search...

771,338 PUBLICATIONS	28,029 DATA SETS	57,558 RESEARCH	49,744 PEOPLE	2,934 ORGANISATIONS	1,733 <small>beta release</small> ENHANCED PUBLICATIONS
-------------------------	---------------------	--------------------	------------------	------------------------	--

SCIENTIFIC NEWS FROM OTHER SOURCES (DUTCH)

- > Internetverslaving niet, eetbuien wel aangemerkt als ziekte
18-05-2013 | Elsevier
- > 'Zorgverzekeraar moet fraude ziekenhuisnota's aanpakken'
18-05-2013 | de Volkskrant
- > Te nerveus voor je sollicitatiegesprek?
18-05-2013 | Intermediair Nieuws
- > More scientific news

Figura 6. Il Portale unificato degli Archivi aperti nazionali olandesi della ricerca NARCIS

II. Modello tecnico che riguarda la raccolta (*harvesting*) di metadati da parte di SP tramite il protocollo di interoperabilità OAI-PMH, attraverso le seguenti richieste:

- a) *GetRecord* = raccogliere *record* di metadati;
- b) *ListIdentifier* e *ListRecord* = raccogliere le liste di identificatori delle unità informative e di record di metadati;
- c) *Identify* = raccogliere le informazioni generali sugli Archivi e sui loro contenuti;
- d) *ListMetadataFormats* = interpretare i formati di metadati raccolti;
- e) *ListSets* = interrogare gli Archivi e le partizioni in cui i medesimi hanno organizzato i dati.

I SP si pongono più vicino all'utente finale offrendone ricche possibilità di ricerca sui contenuti digitali aggregati. Il contenuto digitale può essere considerato su tre livelli:

- 1) *risorsa* (dati, documenti o oggetti digitali) contenuta negli Archivi mantenuti dai DP e descritta da *record* di metadati (scheda descrittiva);
- 2) *record* che rappresentano i metadati espressi in sintassi XML secondo lo schema di metadati DC *simple*, ma anche secondo gli altri metadati definiti in specifici domini di applicazione per la descrizione delle *Risorse*. Ogni *record* deve avere:
 - a) *record* di metadati (DC *simple*) che lo descrivono (catalogano);
 - b) un codice univoco ed identificativo;
 - c) la struttura secondo un determinato *layout*, perché ogni *Risorsa* è granulare (atomi di informazione collegati, espressi tramite lo schema standard XML);
- 3) *item* è un contenitore di tipo logico a partire dal quale vengono diffusi i metadati.

Il protocollo OAI-PMH - che permette a SP di raccogliere i *record* di metadati da DP - può essere definito come un insieme delle strategie di marketing per DP, fondate sulla tecnologia per l'interoperabilità tecnica delle informazioni a livello di *record*, *item*, e *risorsa* attraverso l'utilizzo di standard tecnici aperti comuni.

Grazie all'OAI-PMH, gli Archivi aperti (DP e SP) possono dialogare non solo tra di loro, ma anche con i principali motori di ricerca - sia specialistici che generalisti (*Google*, *Yahoo*, *Google Scholar*, *Scientific Commons*, *Base*, *OAIster*) - che indicizzano i metadati e li rendono disponibili per la meta-ricerca.

Accanto all'interoperabilità tecnica, è opportuno che per ogni Archivio preposto alla ricerca dei contenuti vengano sviluppati i meccanismi a supporto dell'interoperabilità semantica [V.

Appendice I, “Linguaggi e soluzioni tecnologiche del Web Semantico”] facilitata dall’uso di terminologie e vocabolari controllati.

L’interoperabilità è un aspetto di cui ogni Istituzione, che produce e gestisce i contenuti digitali deve tener conto, in quanto essa:

- facilita il ritrovamento dei contenuti depositati in un Archivio digitali da un ampio numero di agenti (Portali, Aggregatori di metadati, Motori di ricerca);
- produce traffico di contenuti verso l’Archivio di riferimento;
- accresce la visibilità dei contenuti dell’Istituzione.

2.2. Strategia dell’*Open Access* “*Green Road*”: Archivi aperti *Open Access*

Viene definita “*Green Road*” o “*via verde*” - il primo canale dell’*Open Access* per disseminare i metadati e *full-text* delle risorse di ricerca attraverso gli Archivi aperti, appoggiati sul protocollo dell’interoperabilità OAI-PMH.

La strategia “*Green Road*” si realizza tramite la pratica del deposito - chiamata *auto-archiviazione* (“*self-archiving*”) [V. sito: “Guida pratica all’auto-archiviazione”] - dei prodotti della ricerca in un servizio Web pubblicamente accessibile (*Open Access*) in accordo con gli standard internazionali dell’*Open Archive Initiative* e le strategie *Open Access* di cui nel Capitolo precedente. L’operazione di *auto-archiviazione* implica l’introduzione nell’Archivio di alcuni dati bibliografici (metadati, indici descrittivi) e l’*upload* (caricamento) del testo completo (*full-text*).

Gli Archivi aperti - secondo l’accezione *Open Access* - comprendono gli Archivi:

1. *Disciplinari* come, ad esempio
 - a) *ArXiv* [V. sito] che contiene in larga parte *pre-prints* di più di mezzo milione di articoli dei fisici e che asseconda un’impostazione di lavoro e di condivisione della ricerca del settore ormai consolidata;
 - b) *RePEc* (*Research Papers in Economics*) [V. sito], la più grande collezione decentrata di documenti ad Accesso Libero per l’economia, proveniente da una rete che collega diversi Archivi) [De Robbio, 2007, p. 87-114].
2. *Istituzionali* per promuovere la visibilità e la consultazione dei lavori di ricerca delle Istituzioni accademiche ed Enti di ricerca.

Elenchi di Archivi aperti *Open Access* si possono consultare in:

- 1) OpenDOAR - *Directory of Open Access Repositories* [V. sito] - un repertorio di oltre 2000 Archivi aperti *Open Access* ricercabile per area disciplinare (tenendo anche conto del fatto che molti Archivi Istituzionali rientrano nella categoria “*multidisciplinary*”), lingua, tipo di materiale, nazione, Software utilizzato);
- 2) ROAR – *Registry of Open Access Repositories* [V. sito] - un registro di circa 2000 Archivi aperti ricercabili *Open Access* per: nazione, tipo di Archivio o Software utilizzato;
- 3) PLEIADI - Portale per la letteratura scientifica elettronica italiana su Archivi Aperti e Depositi Istituzionali [V. sito].

È in corso dal 2008 (in due edizioni annuali) l'iniziativa del CSIS (*Consejo Superior de Investigaciones Científicas in Spagna*) "Ranking Web of Repositories" [V. sito] (Figura 7), che classifica più di 1200 Archivi aperti di tutto il mondo. Il ranking avviene secondo una combinazione dei seguenti criteri:

1. *Dimensione*,
2. *Visibilità*;
3. *Rich Files* (.pdf estratti da Google);
4. *Scholar* (la media tra il numero totale di oggetti rilevati tramite *Google Scholar*; tra il 2007 e il 2012 – per il 2013).



Figura 7. Il servizio "Ranking Web of Repositories"

Un Archivio aperto, anche se nasce come strumento per la disseminazione dei lavori di ricerca, può essere usato per il deposito dei *materiali didattici* per sostenere gli studiosi nel campo dell'insegnamento e dell'apprendimento:

La ricerca è (dovrebbe essere) unita alla didattica e risultati di una ricerca in corso (e tanto più se conclusa) sfociano spesso in una dispensa o in un testo nel quale i concetti della ricerca vengono proposti in un linguaggio didattico. Per questa ragione è importante che i *materiali didattici* guardino verso l'*Open Access* e siano realizzati secondo questa concezione [Guerrini, 2010, p.29].

Se, ad esempio, libri di testo e materiale didattico (e, possibilmente, altri strumenti *online* come ad es. *Blog*) sono stati legati alla didattica entro un'Istituzione, questi - tramite la loro descrizione e deposito nell'Archivio Istituzionale - potrebbero essere utilizzati non solo a livello locale ma anche diffusi in tutto il mondo [V. siti: "MIT OpenCourseWare"; "Archivio della didattica"; "ARIEL"].

Alcune Università sviluppano l'Archivio esclusivamente nell'ottica della *ricerca*, altre estendono la definizione del servizio a tutti i materiali prodotti come risultato delle attività istituzionali di *ricerca e didattica*; altre ancora preferiscono avere più *installazioni distinte* (per es. una dedicata alla ricerca, una dedicata alla didattica) [V. sito: "Linee Guida per gli Archivi Istituzionali", p. 10].

Inoltre,

Il vivace dibattito intorno alla realizzazione degli *Open Archive* dovrebbe portare all'avvio di una più forte sinergia tra il mondo della ricerca accademica e quello delle biblioteche [...] Gli *Open Archive* potranno infatti dispiegare le proprie enormi potenzialità solo attraverso il rafforzamento di queste sinergie, salvaguardandole da possibili contrasti settoriali tra bibliotecari, informatici e accademici [Guerrini, 2010, p.147].

Intanto, prima di iniziare la programmazione di un Archivio aperto,

occorre comprendere bene quale dovrà essere il *modello organizzativo* più congeniale [...] Non una questione puramente tecnica perché gli strumenti *Software* per costruire un Archivio ci sono e peraltro sono strumenti facili da utilizzare, da gestire, sono *Open Source*. Serve il *consenso* degli *Organi Accademici*. Servono le *Politiche*, per la *sottomissione dei dati* e dei *metadati*, le politiche di *Copyright* [...] Occorre dargli un'*identità* precisa, scegliere un nome, stabilire *chi può accedere*, *chi può depositare*, *che cosa va depositato* e cosa no, quali sono i *formati accettati*. È necessario [...] stabilire se deve essere un Deposito che contenga "tutto" quello che viene prodotto dall'Ente, *pre-print*, report tecnici, capitoli di libri, interventi a seminari, presentazioni ... o solo ciò che viene sottomesso a periodici o presentato a convegni [De Robbio, 2007, pp. 89-106].

Negli Archivi aperti del materiale di ricerca possono essere *auto-archiviati*:

- 1) i lavori “*pre-print*” (versioni prima della *peer-review*). In questo caso, la conoscenza scientifica circola più rapidamente, accorciando i tempi di latenza delle pubblicazioni scientifiche presso gli editori, avanzando così la ricerca. L’auto-archiviazione di “*pre-print*” è una pratica comune nel già citato Archivio aperto Disciplinare per i fisici ArXiv;
- 2) i lavori “*post-print*” (versioni *peer-reviewed* pubblicate nelle Riviste scientifiche);
- 3) le Tesi di Dottorato;
- 4) la cosiddetta “*grey literature*” (letteratura grigia) che comprende relazioni di convegni, Tesi di Laurea, materiali didattici, rapporti tecnici ecc.

La pratica dell’auto-archiviazione dei materiali di ricerca in Archivi aperti offre i seguenti vantaggi:

- il Libero Accesso in rete ai *metadati* (indici descrittivi) dei lavori depositati;
- il Libero Accesso al *full-text*, in rispetto alle condizioni contrattuali offerte dall’editore e tramite l’applicazione di Licenze Aperte.

Ormai diversi grandi editori internazionali, tra i quali ad esempio *Elsevier*, stanno cominciando ad aprirsi ai nuovi *modelli editoriali in Accesso Aperto*, diventando tendenzialmente favorevoli a consentire agli autori l’auto-archiviazione dei loro articoli di ricerca negli Archivi Istituzionali. *Elsevier* richiede, ad esempio, che venga inserito il *link* alla rivista da cui gli articoli sono tratti o che non si utilizzi il file in “.pdf” dell’editore. Così facendo, l’editore si garantisce ed anzi aumenta la propria visibilità editoriale e commerciale. Dal sito web della banca dati SHERPA [V. sito] si possono trarre informazioni sulle Politiche perseguite da un numero elevato di editori internazionali in materia di consenso al deposito ad Accesso Aperto e di Politiche editoriali di *Copyright*.

Molti Atenei ed Enti finanziatori di ricerca hanno adottato le Politiche “mandatarie” (*Mandate*, di obbligo) per il deposito dei materiali scientifici prodotti dai ricercatori negli Archivi aperti *Open Access*. Dal 2003 fino a 2013 il numero dei mandati *Open Access* adottati da diverse Istituzioni a livello globale è arrivato quasi a 300 (Figura 8). Nel corso del 7° Programma Quadro (2007-2013) l’Unione Europea ha avviato un Progetto Pilota *Open Access* “*OpenAIRE*” [V. sito] per sostenere la creazione di una rete di Archivi ad Accesso Aperto per il deposito obbligatorio dei materiali che riportano le ricerche finanziate con i fondi pubblici in sei aree disciplinari:

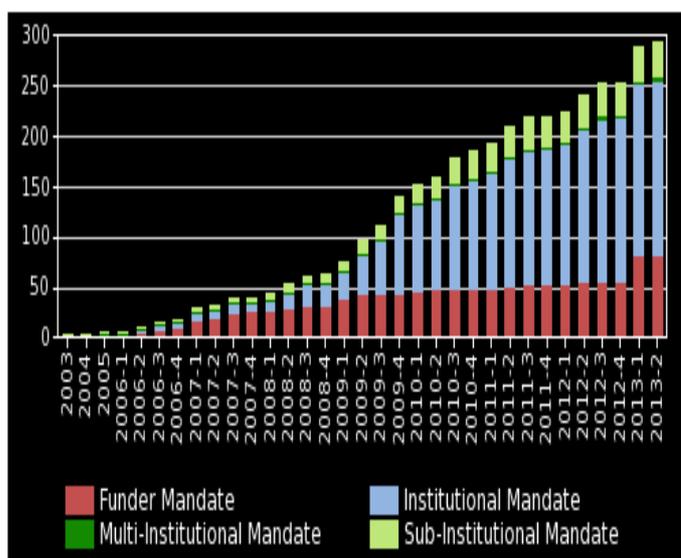


Figura 8. Il numero dei mandati *Open Access* adottati da diverse Istituzioni a livello globale (2003- 2013)

Energia, Ambiente, Salute, ICT, Scienze Sociali, Scienze Umane. L’elenco di centri e fondazioni di ricerca che richiedono ai loro ricercatori di fornire gli articoli “*peer-reviewed*” tramite l’auto-archiviazione in Archivi aperti *Open Access* si può consultare in *Registry of Open Access Mandatory Archiving Policies (ROARMAP)* [V. sito].

Il documento “*The Institutional Repository: Benefits and Challenges*” (*American Library Association, 2013*) [V. sito] fornisce i consigli pratici su come definire una Politica istituzionale per la promozione dell’Accesso Aperto e come implementare un Archivio aperto *Open Access*.

Il documento *"Incentives, Integration, and Mediation: Sustainable Practices for Populating Repositories"* [V. sito] - pubblicato da COAR (*Confederation of Open Access Repositories*) nel 2013 – descrive le pratiche su come popolare i *Repository Open Access*.

Prima di concludere, occorre ribadire che molti Archivi aperti *Open Access*, accanto all'obiettivo principale di promuovere l'auto-archiviazione e la libera disseminazione e condivisione della conoscenza sui lavori del sapere scientifico su larga scala (contribuendo così al progresso della ricerca scientifica), offrono anche diversi servizi a valore aggiunto, che permettono di:

- a) *tracciare* le statistiche d'uso (*visite*, accesso e *download*) del materiale depositato;
- b) *scoprire* gli *interessi di ricerca* di un autore o gruppo di ricerca (classificando e attribuendo *ranking* ai contributi);
- c) *eseguire la ricerca semantica dei contenuti* attraverso le voci di Thesaurus implementati
- d) rilevare i *contenuti correlati* entro e fuori le proprie piattaforme di ricerca [V. siti: OpenAGRIS; VOA3R];

e altro ancora [V. sito: "Funzionalità offerte dall'Archivio Istituzionale dell'Università di Padova PHAIDRA"].

Ipotizzando che tutti gli Archivi aperti *Open Access* fossero organizzati in un'unica rete mondiale interoperabile, il vantaggio per tutta la comunità di ricerca diverrebbe inimmaginabile.

2.2.1. Archivi aperti Istituzionali *Open Access*: scopi, funzionalità

Gli Archivi Istituzionali possono essere considerati degli indicatori tangibili della *qualità di una istituzione accademica*, la sua estensione naturale in quanto volani della ricerca primaria – potenzialmente la componente più importante nell'evoluzione dei nuovi modelli di comunicazione scientifica [V. sito: "Linee guida per gli Archivi Istituzionali", p. 6].

L'idea dell'Archivio Istituzionale risale all'anno 1994 (fino a quel momento esisteva solo la concezione di raggruppamento dei documenti di singole discipline), quando Stevan Harnad (uno tra i primi sostenitori *Open Access*) tramite la *mailing list* del *Virginia Polytechnic Institute* propose di condividere i lavori di ricerca delle Istituzioni accademiche tramite Internet, per renderli più fruibili e accessibile in rete dalla comunità degli studiosi e da altri interessati, in modo che i lavori di ricerca potessero essere valutati e commentati dai pari [Poynder, 2004].

Con le successive iniziative del movimento internazionale *Open Access* – entro cui è stato formalizzato il concetto degli Archivi aperti Istituzionali compresi nella strategia "*Green Road*" - la comunicazione dell'informazione scientifica è cambiata profondamente:

l'articolo pubblicato su una Rivista rappresenta spesso il punto finale di una catena comunicativa al cui inizio c'è un intervento a un convegno. Gli Archivi ad Accesso Aperto permettono di valorizzare tutte le tappe di questa catena e di far crescere la reputazione dell'autore e il futuro impatto dell'articolo. È stato anche dimostrato che gli articoli depositati in Archivi ad Accesso Aperto sono citati più spesso di quelli che non lo sono

[V. sito: "*Open Access* e la valutazione" CRUI, p.5].

Un Archivio Istituzionale (*Institutional Repository, IR*) *Open Access* è un Archivio aperto con l'obiettivo di favorire il deposito (auto-archiviazione), l'accesso aperto, la distribuzione illimitata, la visibilità, l'interoperabilità e l'impatto dei risultati dell'attività di ricerca dei ricercatori di un Ente a livello globale, con il supporto di tecnologie di rete.

L'aggettivo "*Istituzionale*" fa riferimento a "un ruolo ben definito nella comunicazione scientifica e ad una chiara *identità istituzionale*, caratteristica che rende [Archivio Istituzionale] diverso da un Archivio Disciplinare o "*subject-based repository*" [Guerrini, 2010, p.25].

In alcuni casi potrebbe essere previsto che gli autori depositino (auto-archiviano) i loro risultati di ricerca in più *Repository* contemporaneamente ad esempio:

- a) nell'Archivio dell'Istituzione di appartenenza;
- b) negli Archivi di co-autori;
- c) negli Archivi Disciplinari;
- d) negli Archivi di finanziatori di ricerca;

allo scopo di sostenere uno scenario comune per la ricerca scientifica.

In questa prospettiva, tra i principali obiettivi dell'Archivio Istituzionale deve inserirsi anche il supporto al processo del “*deposito multiplo*”, tramite l'applicazione degli appositi strumenti per il trasferimento dei contenuti da un sistema ad un altro.

Tra gli strumenti che appoggiano il processo di trasferimento dei contenuti di ricerca negli Archivi aperti multipli, anche in collegamento con i sistemi informativi della ricerca *Current Research Information Systems* (CRIS; EuroCRIS: un approccio comune per l'amministrazione della ricerca in Europa [V. siti]),

è opportuno citare i protocolli:

- a) SWORD (*Simple Web-service Offering Repository Deposit*), che è stato implementato nelle versioni recenti delle piattaforme *Software* (DSpace, EPrints, Fedora, Invenio, Zentity, Intralibrary) per l'implementazione di Archivi aperti [V. siti: “ArXiv - SWORD/APP Deposit”; SWORD];
- b) OA-RJ (*Open Access Repository Junction*).

I medesimi protocolli consentono facilmente di depositare i contenuti digitali agendo da una postazione remota verso un Archivio Aperto e viceversa, e anche da un Archivio verso più Archivi, tramite lo scambio metadati tra diverse piattaforme. Ambedue i protocolli – assieme con altri modelli per l'interoperabilità - sono al centro del Progetto del *Knowledge Exchange “CRIS-OAR”* (*Current Research Information Systems - Open Access Repositories*) [V. sito: “*Current State of Open Access Repository Interoperability*”].

Le “*Linee guida per gli Archivi Istituzionali*” CRUI (2009) - che forniscono le *linee di indirizzo comuni* per la creazione degli Archivi Istituzionali - definiscono questi ultimi

come uno strumento concreto a disposizione della comunità scientifica, un'infrastruttura informativa e comunicativa che raccoglie in *un unico luogo* tutta la produzione scientifica di un Ateneo, alla quale viene conferita *autorevolezza*, assicurandone nel contempo la *persistenza in rete* e la conservazione a lungo termine [V. sito: “*Linee guida per gli Archivi Istituzionali*”, p.6].

Il sostenitore dell'*Open Access* - Guédon definisce gli Archivi Istituzionali come il modello ‘*Overlay Journals*’ definito anche ‘*Article Database*’ e ‘*Deconstructed Journal*’ o meglio le riviste che non producono un proprio contenuto, ma selezionano testi dagli autori.

A titolo esemplificativo, si cita il “*Bollettino telematico di Filosofia Politica*” [V. sito] - pubblicato dalla Facoltà di Scienze Politiche dell'Università di Pisa - che è diventato un *Overlay Journal*, una rivista “copertina” *peer-reviewed*, che recensisce e segnala articoli ad Accesso Aperto depositati altrove.

Occorre comunque tener presente che un Archivio aperto per la disseminazione dei materiali di ricerca

non è un database catalografico, ha scopi e obiettivi *diversi* rispetto ad un *Catalogo*, in quanto si fonda e nasce su presupposti diversi. E non è nemmeno un *Catalogo* nel senso più profondo del termine, non è un *simil-MARC* per intenderci; i suoi metadati sono di basso livello, nella maggior parte degli Archivi si usa *Dublin Core* non qualificato [...] o al massimo *Dublin Core* con qualificatori, come nel caso delle esperienze con la piattaforma *DSpace*. Per tale ragione strumenti per l'*Authority file*, *Thesauri*, o *Schemi di Classificazione* sono demandati ai *Service Provider* che potremmo definire come servizi a valore aggiunto costruiti attorno, dietro, davanti, in mezzo, sopra i *Data Provider*, ad un livello meta di aggregazione [De Robbio, 2007, p.250-251].

Gli Archivi Istituzionali vengono sviluppati per raccogliere e disseminare in rete (“*one input, many outputs*”) tutta la testimonianza della produzione intellettuale dell'ente, coinvolgendo i contenuti digitali di varie discipline registrati attraverso:

- 1) *Articoli scientifici*;
- 2) *Tesi di Dottorato*;
- 3) *Dati di ricerca (Research Data)* [V. sito];
- 4) “*Letteratura grigia*” (atti di convegni, brevetti, documenti multimediali ecc.).

Nelle Università italiane, entro il 2013, sono stati implementati 47 Archivi Istituzionali [V. sito: “PLEADI”: Archivi Istituzionali]:

1. ACUBO, Università Politecnica delle Marche - <<http://openarchive.univpm.it/>>
2. AISBERG, Università degli studi di Bergamo - <<http://aisberg.unibg.it/>>
3. AIR, Università degli studi di Milano - <<http://air.unimi.it/>>
4. AMS ACTA, Università degli studi di Bologna - <<http://amsacta.cib.unibo.it/>>
5. AMS Dottorato, Università degli studi di Bologna - <<http://amsdottorato.cib.unibo.it/>>
6. AMS Laurea, Università degli studi di Bologna - <<http://amslaurea.cib.unibo.it/>>
7. APERTO, Università degli studi di Torino - <<http://aperto.unito.it/>>
8. ARCADIA, Università degli studi Roma Tre - <<http://dspace-roma3.caspur.it/>>
9. Archenvimat, CNR - <<http://www.archenvimat.pz.cnr.it/>>
10. ARchivIA, Università degli studi di Catania - <<http://dspace.unict.it/>>
11. ART, Università degli studi di Roma Tor Vergata - <<http://art.torvergata.it/>>
12. BESS Digital Archive, Biblioteca Elettronica di Scienze Sociali ed Economiche del Piemonte - <<http://www.digibess.it/>>
13. BOA, Università degli studi di Milano-Bicocca - <<http://boa.unimib.it/>>
14. CADMUS, Istituto Universitario Europeo - <<http://cadmus.eui.eu/>>
15. CAMEPRINTS, Università degli studi di Camerino - <<http://cameprints.unicam.it/>>
16. CEUM, Università degli studi di Macerata - <<http://archiviodigitale.unimc.it/>>
17. DOCTA, Università Cattolica - <<http://tesionline.unicatt.it/>>
18. DSpace ISS, Istituto Superiore di Sanità - <<http://dspace.iss.it/>>
19. DSpace@UniPR, Università degli studi di Parma - <<http://dspace-unipr.cilea.it/>>
20. Earth-Prints, INGV - <<http://www.earth-prints.org/>>
21. EleA, Università degli studi di Salerno - <<http://elea.unisa.it:8080/jspui/>>
22. E-STOA, STOA - <<http://eprints.stoa.it/>>
23. E-LIS, Eprints in Library and Information Science - <<http://eprints.rclis.org/>>
24. FEDOA, Università degli studi di Napoli Federico II - <<http://www.fedoa.unina.it/>>
25. InsubriaSpace, Università degli studi dell'Insubria - <<http://insubriaspace.cilea.it/>>
26. LEAR, Università Ca' Foscari di Venezia - <<http://lear.unive.it/>>
27. LUISSearch, LUISS - <<http://eprints.luiss.it/>>
28. MoReThesis, Università degli studi di Modena e Reggio Emilia - <<https://morethesis.unimore.it/>>
29. PADIS, La Sapienza Università di Roma - <<http://padis.uniroma1.it/>>
30. Padua@Research, Università degli studi di Padova - <<http://paduaresearch.cab.unipd.it/>>
31. Padua@Thesis, Università degli studi di Padova - <<http://tesi.cab.unipd.it/>>
32. PORTO, Politecnico di Torino - <<http://porto.polito.it/>>
33. OPAR, Università degli studi di Napoli L'Orientale - <<http://opar.unior.it/>>
34. OpenPub, IASMA - <<http://open-pub.iasma.it/>>
35. OpenStarts, Università degli studi di Trieste - <<http://www.openstarts.units.it/>>
36. OpenUniud, Università degli studi di Udine - <<http://dspace-uniud.cilea.it/>>
37. ROAD, Università degli studi del Molise - <<http://road.unimol.it/>>
38. SISSA Digital Library, SISSA - <<https://digitallibrary.sissa.it/>>
39. SOLAR, CNR - <<http://eprints.bice.rm.cnr.it/>>
40. SURplus UniPA, Università degli studi di Palermo - <<http://surplus.unipa.it/>>
41. UniCa Eprints, Università degli studi di Cagliari - <<http://veprints.unica.it/>>

42. UniFI-UGOV, Università degli studi di Firenze - <<http://www.unifi.u-gov.it/>>
43. UnissResearch, Università degli studi di Sassari - <<http://eprints.uniss.it/>>
44. UNITN-eprints, Università degli studi di Trento - <<http://eprints.biblio.unitn.it/>>
45. UNITN-eprints PhD - <<http://eprints-phd.biblio.unitn.it/>>
46. UniTus DSpace, Università degli studi della Toscana - <<http://dspace.unitus.it/>>
47. UniVR-UGOV, Università degli studi di Verona - <<http://www.univr.u-gov.it/>>

La maggior parte di questi Archivi [V. sito: “*Wiki.Openarchives.it: Applicazione delle Linee Guida*”] sono implementati nel rispetto di Regolamenti dei propri Atenei, relativi al deposito obbligatorio (*mandate*) delle Tesi di Dottorato negli Archivi Istituzionali ad Accesso Aperto. Di seguito si presenta un estratto del Regolamento del bando di concorso per l'ammissione ai corsi di Dottorato di ricerca che dichiara l'obbligo del deposito delle Tesi di Dottorato nell'Archivio Istituzionale *Open Access*:

Università degli studi di Napoli Federico II

Estratto dal bando di concorso per l'ammissione ai corsi di Dottorato di ricerca XXIV ciclo

"ARTICOLO 13 (Titolo di dottore di ricerca e tesi di dottorato)... L'Università archiverà e renderà consultabile in rete il TESTO completo della TESI di dottorato attraverso l'*Archivio Istituzionale* ad ACCESSO APERTO. Potrà essere concesso un periodo, comunque non superiore a tre anni, su motivata richiesta del dottore di ricerca, in cui la tesi non sarà consultabile da parte di terzi."

Per supportare il processo di deposito delle tesi del dottorato, il Gruppo *Open Access* della CRUI ha messo a disposizione delle Università Italiane le “Linee guida per il deposito delle Tesi di Dottorato negli Archivi aperti” (2007) [V. sito], “prendono le mosse dal riconoscimento delle *Tesi di Dottorato* come *prodotti della ricerca* a tutti gli effetti, che quindi, secondo la Dichiarazione di Berlino e le recenti raccomandazioni della Commissione Europea, dovrebbero essere *pubblicamente accessibili*” [ivi, p.3], sulle quali la comunità scientifica potrebbe lavorare e discutere.

Le Università e i Centri di ricerca, *piuttosto che preoccuparsi della segretezza e confidenzialità di dati sensibili* e della mancata divulgazione dei risultati delle ricerche *per scopi non sempre trasparenti*, dovrebbero assicurarsi che i *risultati della ricerca* che loro stessi finanziano e appoggiano in tutti i modi, *abbiano il più alto livello di impatto entro le comunità* [De Robbio, 2007, p. 42].

Le caratteristiche e le funzionalità principali di un Archivio aperto Istituzionale di ricerca possono essere riassunte tramite i seguenti punti (Tabella 1).

Caratteristiche	Funzionalità
principali	
1. Il materiale è di natura scientifica 2. Il materiale è creato presso l'Istituzione ospitante 3. Vengono utilizzate le tecniche di <i>auto-archiviazione</i> (deposito) del materiale 4. L'Accesso aperto immediato al maggior numero dei contenuti depositati in <i>full-text</i>	1. Applicazione di metadati per la descrizione e strutturazione dei contenuti digitali 2. Controllo degli accessi 3. Supporto alla scoperta delle risorse 4. Distribuzione del materiale in rete 6. Preservazione (conservazione digitale a lungo termine) e l'accesso persistente [Pasqui, 2012].
Interoperabilità tecnica tramite il protocollo OAI-PMH: un Archivio aperto Istituzionale dialoga con gli altri Archivi sviluppati sulle piattaforme <i>Open Archive Initiative</i> e con i principali motori di ricerca (sia specialistici che generalisti) che indicizzano i metadati dell'Archivio e li rendono disponibili per la meta-ricerca.	

Tabella 1. Caratteristiche e funzionalità principali di un Archivio Istituzionale

Considerando che gli Archivi Istituzionali devono fungere “da indicatori tangibili della qualità della ricerca di un’Istituzione aumentandone la visibilità, il prestigio e il valore pubblico” [V. sito: “*Open Access e la valutazione*”, p.14], ogni Archivio deve implementare i *sistemi per verifica di qualità* di:

- a) materiali depositati;
- b) *metadattazione* (indicizzazione) descrittiva e semantica a supporto dei materiali depositati

[V. sito: *Metadata & Semantics for Open Access Repositories*].

Ciò inciderà positivamente sul livello di qualità dei contenuti depositati, sul reperimento dell'informazione, sull'efficacia dell'*harvesting* (raccolta) di metadati operato dai SP e motori di ricerca (*Google, Google Scholar*), e sulla valutazione della ricerca dell'Istituzione.

2.2.1.1. Linee di indirizzo a supporto dell'implementazione e della gestione

Prima di stabilire il modello organizzativo, o meglio, la tipologia dell'Archivio Istituzionale che si vuole offrire alla propria comunità (l'Archivio del materiale di ricerca; di ricerca e didattica; Archivi con installazioni distinte), devono essere ben definiti:

- 1) fabbisogni dei potenziali utenti istituzionali;
- 2) finalità ci si prefigge raggiungere attraverso il medesimo Archivio, più convenienti alle specifiche situazioni;
- 3) funzioni dell'Archivio.

Una chiara identificazione di questi punti implica una serie di *scelte politiche* (criteri, regole, responsabilità e ruoli di organizzazione) e *organizzative* (flussi di lavoro, unità/sistemi coinvolti, schemi di metadati, formati, diritti di proprietà intellettuale ecc.) che devono essere dichiarate attraverso le *linee d'indirizzo* dell'Archivio Istituzionale, che possano guidare lo sviluppo del medesimo Archivio e aiutare a risolvere una serie di problemi che i gestori dell'Archivio possono incontrare.

Le *linee d'indirizzo* dovrebbero inoltre includere le informazioni relative a:

- 4) la descrizione dell'organizzazione dell'architettura dell'informazione, particolarmente:
 - a) il disegno dello schema dei metadati a supporto dei materiali da depositare e dell'interoperabilità dei contenuti attraverso i metadati:
 è importante che si definiscano con chiarezza le *tipologie di materiale e i metadati* che verranno utilizzati per descriverle [...] adattando quanto previsto dalle norme di catalogazione alle esigenze

della valutazione e procedendo, dove è il caso, anche a delle integrazioni. L'interoperabilità così raggiunta consentirebbe il colloquio non solo con l'eventuale Archivio nazionale della ricerca, ma anche con i motori di ricerca specializzati (è il caso, ad esempio, di Wageningen Yeld, l'Archivio dell'Università di Wageningen interrogabile anche tramite *Google Scholar* e *Scirus*) [V. sito "Open Access e la valutazione", p.14].

A proposito della scelta di metadati per la descrizione dei contenuti di ricerca è utile riferirsi alle "Linee guida per la creazione e la gestione di metadati nei Repository Istituzionali" CRUI (2012) [V. sito];

- b) collegamenti tra i materiali (es. il collegamento dell'Archivio/della Collezione delle Tesi di Dottorato con l'Archivio delle pubblicazioni e così via);
- c) il controllo della qualità di metadati attraverso la loro bonifica con *file* di autorità (vocabolari controllati, Thesauri) da inserire nel sistema;

5) l'elaborazione di una "guida" alla *proprietà intellettuale* attivata sia a livello di processo dell'auto-archiviazione, sia a livello di gestione delle problematiche di natura corrente tramite l'attivazione di un *Help Desk* di supporto.

Basandosi sulle *Linee d'indirizzo*, si potrebbe partire con l'implementazione dell'*Archivio pilota* tramite la sua sperimentazione entro una *Comunità pilota* (un Dipartimento, un Centro di ricerca; in generale una o più "Comunità di interesse"), i membri della quale potrebbero iniziare a depositare nel medesimo Archivio la propria produzione intellettuale di ricerca, "allo scopo di testare le procedure e di avere un riscontro concreto in termini di "gradimento", nonché un supporto nella redazione di norme e buone pratiche finalizzate al "popolamento" dell'Archivio" [V. sito: "Linee Guida per gli Archivi Istituzionali", p. 11]. Al fine di identificare gli scenari di crescita dell'*Archivio Pilota*, è opportuno revisionare periodicamente il suo modello sperimentale, in particolar modo il grado di maturità del servizio in rapporto alle finalità attese, valutando le funzionalità effettive e monitorando il rendimento dell'Archivio. Quando l'Archivio inizia a 'radicarsi' nell'ente, entrando in una fase operativa vera e propria, sarebbe opportuno iniziare a definire gli aspetti ritenuti prioritari per la strutturazione dell'*Archivio* come *servizio*. In questa fase le decisioni possono prendere la forma di una *Politica (Policy)*, che è sempre più flessibile rispetto a quella di Statuto o Regolamento. Secondo le citate "Linee Guida per gli Archivi Istituzionali", entro la *Politica* dell'Archivio Istituzionale devono essere definiti i seguenti punti (Tabella 2):

1.	Obiettivi e funzioni a supporto dell'Archivio
2.	Contenuti (tipologie di materiali; formati accettati)
3.	Organizzazione dei materiali
4.	Persone autorizzate (o tenute) all'auto-archiviazione dei materiali
5.	Ruoli e le responsabilità coinvolte nel flusso di lavoro
6.	Aspetti legali: le responsabilità degli autori rispetto agli eventuali vincoli contrattuali con gli editori
7.	Modalità di accesso
8.	Standard di metadati
9.	Procedure per l'inserimento e la verifica di metadati
10.	Verifica qualitativa di materiali
11.	Formati per assicurare l'integrità dei materiali ospitati nel tempo (preservazione)
12.	Garanzie di sicurezza e procedure di <i>back-up</i>
13.	Circostanze e le relative procedure per un'eventuale rimozione di materiali depositati
14.	Modalità di gestione di dati personali (<i>Privacy</i>)
15.	Modalità di monitoraggio dell'Archivio per una migliore divulgazione del materiale depositato

Tabella 2. Punti principali da considerare in una *Politica (Policy)* dell'Archivio Istituzionale

C'è necessità di una *Policy* chiara e condivisa rispetto agli oggetti da depositare e ai metadati necessari a descriverli, ai soggetti titolati a depositare i propri lavori, alla tempistica entro cui depositare, all'*obbligo di deposito* almeno dei *metadati* e preferibilmente anche dei *full-text*, alla definizione di un *workflow* che preveda la *validazione dei metadati* prima della loro pubblicazione [Galimberti, 2010, p.31].

Alcune caratteristiche presentate nella Tabella 2 possono variare in base agli obiettivi, alle dimensioni e alle disponibilità delle risorse dell'Ente che intende adottare una *Politica (Policy)* dell'Archivio Istituzionale.

L'auto-archiviazione dei materiali di ricerca nell'Archivio può avvenire a cura degli autori (nella maggior parte dei casi), o da terze persone (bibliotecari).

In fase operativa vera e propria di un Archivio Istituzionale, l'auto-archiviazione può definita essere volontaria o mandataria con un'eventuale formulazione della *Mandatory Policy* (Figura 9). A

ben vedere, questo tipo di regolamento [*mandate*] *non fissa un vero e proprio obbligo* per l'autore, piuttosto definisce un onere cui questi deve conformarsi se vuole ottenere un beneficio (Accesso a un finanziamento; ammissione alla discussione di una tesi per il completamento di un ciclo di studi svolto avvalendosi di strutture e risorse dell'ente) [Guerrini, 2010, p.123].

Una valida strategia alternativa per ottenere un numero elevato di contributi è quella di introdurre *forme di incentivo che possano stimolare l'interesse degli autori nei confronti dell'Archivio*: per esempio l'erogazione di quote di finanziamento subordinate al deposito di lavori di ricerca, oppure l'attivazione di servizi che possano assicurare agli autori un migliore impatto, un'ampia visibilità e un facile accesso ai loro lavori

[V. sito: "Linee Guida per gli Archivi Istituzionali", pp. 16, 24].

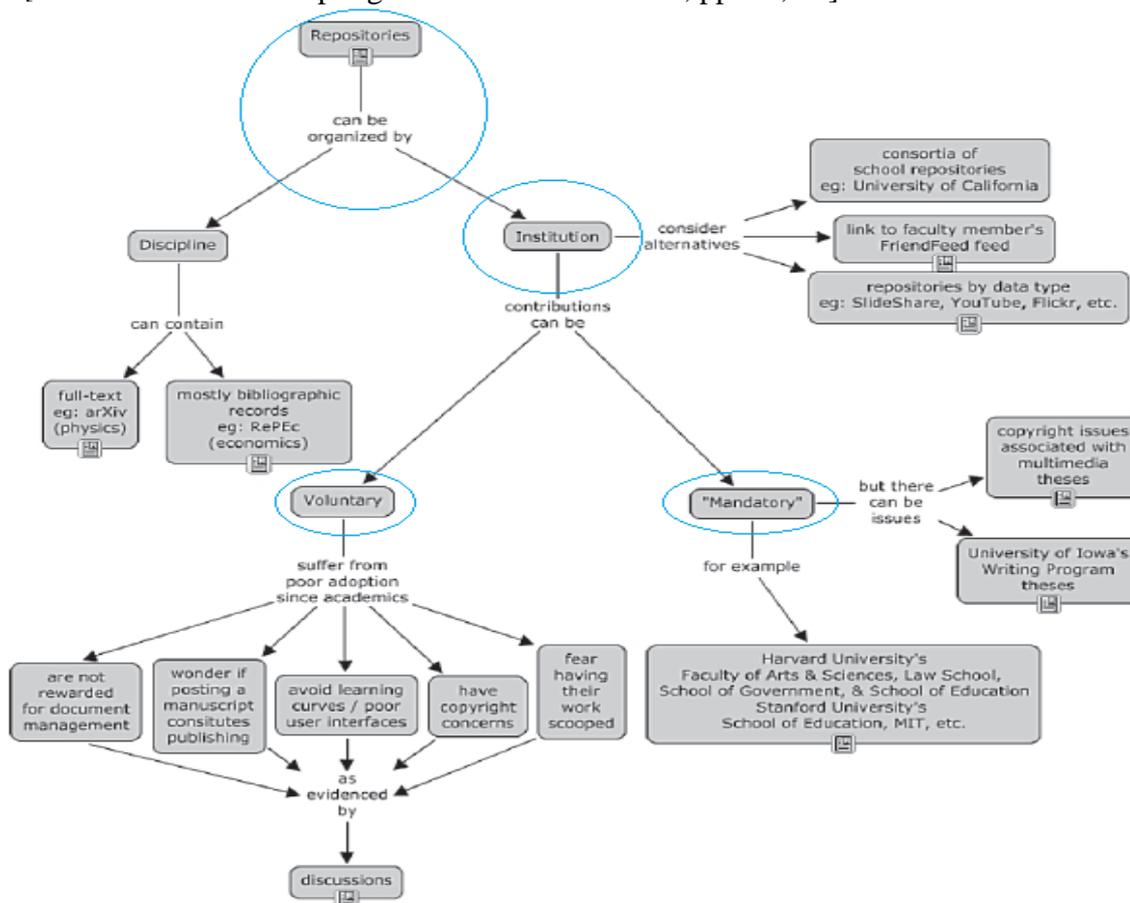


Figura 9. Gli Archivi Istituzionali: le strade per il deposito volontario e mandatario [V. sito "Open Access Information and Resources"]

La definizione di Politiche che prevedano l'obbligo del deposito della produzione scientifica negli Archivi aperti Istituzionali si potrebbe collocare come una potente strategia di rafforzamento della posizione contrattuale dell'autore nel rapporto con gli editori commerciali:

poiché in tal modo egli non appare più portatore di un interesse isolato alla disseminazione senza barriere dei suoi prodotti, ma esprime la posizione di un intero Ateneo o Istituto di ricerca, ovvero di tutti gli autori operanti in quell'Ente. È auspicabile quindi il diffondersi di Politiche istituzionali di questo tipo e, meglio ancora, l'affermazione del principio dell'Accesso Aperto a livello normativo [Guerrini, 2010, p.123].

Riagganciandoci alla questione di programmazione del modello organizzativo di Archivi Istituzionali, occorre dire che essi non vengono progettati al solo scopo di costituire una vetrina "trasparente" per risultati di ricerca dei rispettivi Atenei, bensì anche per contribuire alla realizzazione dei sistemi informativi di questi ultimi; per cui vanno programmati nel contesto più ampio dell'ambiente informativo, integrandosi da un lato con la Biblioteca Digitale,

in modo da consentire una visibilità totale e trasversale da qualsiasi punto del sistema informativo: soluzione Portale, OPAC, Banche Dati, Periodici Elettronici, e in modo da garantire una connessione tra referenze bibliografiche [ivi],

dall'altro con i sistemi amministrativi di Ateneo che gestiscono le pagine *web*, i dati di studenti e docenti, le registrazioni per l'attività didattica, con i settori dell'*e-learning*, con gli Archivi Disciplinari, con le basi di dati per la gestione dell'Anagrafe della ricerca [Mornati, 2006, p.].

Il materiale degli Archivi Istituzionali potrebbe avere inoltre una funzione importante nello sviluppo regionale, permettendo alle aziende, agli Enti pubblici e alle organizzazioni civili di comprendere immediatamente che tipo di *expertise* è disponibile a livello locale.

Gli Archivi Istituzionali possono divenire anche "testimoni della loro *cultura e memoria storica* [delle loro Istituzioni e] strumenti di *valutazione* delle attività e delle *produzioni intellettuali* di ogni singola Istituzione" [Marchitelli, 2007, p. 16].

2.2.1.1.1. Raccomandazioni: il ruolo e alcuni casi d'uso

Nel Diritto Comunitario la "Raccomandazione" è un atto non vincolante a disposizione di varie istituzioni, dirette a sollecitare il destinatario ad adottare un determinato comportamento, o a cessare da un comportamento specifico [V. sito "Portale del Diritto"].

In uno stato le Raccomandazioni sono normalmente dirette dalle sue Istituzioni e contengono l'invito a conformarsi ad un certo comportamento. Sono in genere emanate dalle Istituzioni che non dispongono del potere di emanare atti obbligatori o quando ritengono che non vi sia motivo di ricorrere a norme vincolanti. Sia Raccomandazioni che Linee guida servono per l'orientamento di coloro che collaborano al raggiungimento di questo obiettivo comune. Esse si evolvono secondo le esigenze di un certo dominio applicativo, ma non essendo documenti con le regole fisse, non pretendono di coprire ogni possibile variante di questa o quella questione.

È opportuno che gli Archivi Istituzionali siano istituiti e gestiti in conformità alle migliori pratiche del momento, seguendo, quindi, le specifiche Raccomandazioni, Linee guida e Standard nel settore, affinché possano raggiungere certi livelli di qualità. In questo modo si può contribuire:

- a) all'applicazione delle regole comuni e, perciò, unificanti per la gestione dei contenuti degli Archivi;
- b) alla creazione di una migliore comunicazione e, perciò, all'interoperabilità tra i contenuti di diversi Archivi;
- c) all'aumento della visibilità, l'accessibilità e l'impatto dei contenuti di ricerca depositati negli Archivi.

Nella Tabella 3 vengono elencate alcune Raccomandazioni cui ci si può riferire durante le procedure di pianificazione, implementazione e gestione degli Archivi Istituzionali *Open Access*.

	Raccomandazioni	Riferimenti web
1.	Raccomandazioni per lo sviluppo di Archivi Istituzionali. <i>Enabling Open Scholarship (EOS)</i> (2007)	< http://www.openscholarship.org/jcms/c_6162/en/repositories >
2.	Raccomandazioni del <i>Working Group</i> sull'Accesso Aperto della <i>European University Association (EUA)</i> . Traduzione italiana CRUI (2008)	< http://www.cruir.it/HomePage.aspx?ref=1149# >
3.	Raccomandazioni sull' <i>Open Access</i> e la valutazione dei prodotti della ricerca scientifica CRUI (Conferenza di Rettori delle Università Italiane) (2009)	< http://www.cruir.it/HomePage.aspx?ref=1782 >
4.	Raccomandazioni <i>sull'Open Access</i> per le istituzioni accademiche e di ricerca. EUA (<i>European University Association</i>) (2008)	< http://www.eua.be/fileadmin/user_upload/files/Policy_Positions/Recommendations_Open_Access_adopted_by_the_EUA_Council_on_26th_of_March_2008_final.pdf >
5.	<i>Recommendations of a set of metadata properties and encoding vocabularies. Meaningful Bibliographic Metadata (M2B). AIMS (Agricultural integrated management services)</i> (2012)	http://aims.fao.org/metadata/m2b >
6.	LODE-BD (<i>Linked Open Data – enabled Bibliographical Data</i>) <i>Recommendations (V.2.0)</i> . AIMS (<i>Agricultural integrated management services</i>) (2012)	< http://aims.fao.org/community/metadata/blogs/lode-bd-recommendations-20-are-now-available-html >

Tabella 3. Alcune Raccomandazioni che possono essere applicate per la programmazione e la gestione degli Archivi Istituzionali

Le Raccomandazioni presentano il *know-how* necessario per poter ampliare e ridefinire gli orizzonti dei modelli organizzativi di Archivi Istituzionali esistenti, dando luogo allo sviluppo di una *core set* di caratteristiche tecniche e organizzative per una migliore gestione di contenuti digitali in questi sistemi.

2.2.1.1.2. Linee guida: il ruolo e alcuni casi d'uso

Negli aspetti generali e concettuali di massima, le Linee guida possono essere definite come Raccomandazioni sviluppate attraverso il processo sistematico di elaborazione per assistere i gestori di sistemi (qui Archivi) nelle decisioni da prendere sulla gestione appropriata dei medesimi.

Se nella progettazione e gestione degli Archivi Istituzionali “mancano Linee guida e regole, [essi] incontrano maggiori difficoltà ad acquisire contenuti” [V. sito “Linee guida per gli Archivi Istituzionali”, 2009, p.14].

Le Linee guida si contraddistinguono rispetto a Standard, o protocolli, che presentano strumenti di maggiore rigidità, finalizzati alle azioni ritenute ottimali e quasi obbligate. Una Linea guida,

sostanzialmente, è un risultato di un preciso percorso sistematico di analisi dei processi di un sistema orientato alla definizione della “*Best Practice*”.

L'enorme popolarità che le Linee guida hanno conquistato nei diversi settori può essere spiegata per le funzioni e i ruoli che queste svolgono:

- 1) il primo ruolo è connesso alla funzione di educazione, formazione ed aggiornamento. Il momento di elaborazione di una Linea guida rappresenta un eccezionale momento formativo ed educativo per chi vi partecipa;
- 2) il secondo ruolo è di creare le condizioni per rendere possibile il monitoraggio pratico di un sistema, individuando i processi più appropriati, il loro utilizzo ed i conseguenti risultati ottenuti;
- 3) una terza funzione è la promozione del miglioramento continuo dell'attività assistenziale in quanto le Linee guida sono fondamentali strumenti a base di una azione di governo (*governance*) di un sistema.

Tali passi sono presupposti fondamentali per la realizzazione di un sistema operativamente orientato al cambiamento e al miglioramento continuo.

Uno dei requisiti alla base della qualità delle Linee guida è quello di promuovere le linee di indirizzo fondate scientificamente su Raccomandazioni validate e riproducibili.

In pratica, attraverso le Linee guida si cerca di superare un modello semplicistico di acquisizione dell'informazione adeguandolo e migliorandolo in modo che dal singolo professionista si passi alla collaborazione fra più soggetti od organismi professionali che abbiano le competenze e le risorse professionali per raggiungere lo scopo della migliore interpretazione possibile della letteratura scientifica e delle Buone Pratiche disponibili.

Nella Tabella 4 viene presentato un elenco delle Linee guida sviluppate da organismi professionali per supportare lo sviluppo (incluso la definizione delle Politiche) e la gestione degli Archivi Istituzionali.

	Linee Guida	Riferimenti web
1.	<i>Guidelines for the Creation of Institutional Repositories at Universities and Higher Education Organisations</i> (2007)	< http://works.bepress.com/ir_research/h/18 >; < http://www.sisbi.uba.ar/institucional/proyectos/internacionales/Guidelines_IR_english.pdf >
2.	Linee guida per il deposito delle tesi di dottorato negli Archivi Aperti (CRUI, 2007)	< http://www.crui.it/HomePage.aspx?ref=1149 >
3.	DRIVER (<i>Digital Repository Infrastructure Vision for European Research</i>) Guidelines 2.0. Le Linee Guida per Content Providers (Archivi) – che espongono i loro contenuti attraverso il protocollo di interoperabilità OAI-PMH	< http://www.driver-support.eu/documents/DRIVER_Guidelines_v2_Final_2008-11-13.pdf >
4.	Linee guida per gli Archivi Istituzionali (CRUI, 2009)	< http://www.crui.it/HomePage.aspx?ref=1781 >
5.	<i>OpenAIRE Guidelines 1.1</i> (2010) costruite su DRIVER	< http://www.openaire.eu/en/component/content/article/207 >
6.	<i>OpenAIRE Guidelines 2.0</i> (2012) (aggiornamento di <i>OpenAIRE Guidelines 1.1</i>)	< http://www.openaire.eu/en/component/content/article/427-openaire-releases-version-20-of-the-openaire-guidelines >
7.	Linee guida per la creazione e la gestione di	< http://www.crui.it/HomePage.aspx?

	metadati nei <i>Repository Istituzionali</i> (CRUI, 2012)	ref=2066>
8.	<i>Open Access Guidelines for researchers funded by the ERC (European Research Council)</i> (2012)	< http://erc.europa.eu/sites/default/files/document/file/open_access_policy_researchers_funded_ERC.pdf >
9.	Linee Guida per 'interoperabilità Semantica attraverso i <i>Linked Open Data</i> . Commissione di Coordinamento SPC (Sistema Pubblico di Connettività e Cooperazione) (2012)	< http://www.digitpa.gov.it/sites/default/files/allegati_tec/CdC-SPC-GdL6-InteroperabilitaSemOpenData_v2.0.pdf >

Tabella 4. Alcune Linee Guida a supporto di una migliore programmazione e gestione dei contenuti negli Archivi *Open Access*

Le Linee guida “*OpenAIRE*” sviluppate sulla base delle Linee guida “DRIVER” possono essere particolarmente utili per:

- a) i gestori di Archivi Istituzionali nuovi, per definire le Politiche relative ai formati e ai metadati a supporto dei contenuti digitali,
- b) i gestori di Archivi già avviati per ri-orientare le scelte fatte.

Queste Linee guida forniscono uno strumento per mappare (o tradurre) i metadati di Archivi Istituzionali nel formato *Dublin Core* supportato dal protocollo di interoperabilità tra gli Archivi OAI-PMH.

Tutte le Linee Guida presentate nella Tabella 4 sono considerate tra i fattori che hanno maggiore impatto sulle prospettive di affermazione e di crescita degli Archivi Istituzionali, in quanto:

- ✓ costituiscono il quadro di riferimento per gli aspetti programmatici e organizzativi del progetto e contribuiscono a realizzarne gli obiettivi;
- ✓ permettono di attivare le strategie più efficaci per attrarre contributi e promuovere il cambiamento culturale necessario all'affermazione dell'Archivio;
- ✓ sono di supporto alla promozione dell'Archivio, in quanto definiscono con chiarezza le responsabilità riguardo i contenuti, la gestione e la manutenzione

[V. sito: “Linee guida per gli Archivi Istituzionali, 2009, p.14].

2.2.1.2. Collegamenti con *Current Research Information Systems* (CRIS) per la valutazione della ricerca

Ultimamente si è cominciato a vedere gli Archivi Istituzionali come contenitori di una grande quantità di dati non solo bibliografici, e ad utilizzarli come nucleo centrale nei sistemi di CRIS (*Current Research Information System*). Questi sistemi integrano dati provenienti da fonti diverse (*progetti, finanziamenti, anagrafiche, repertori, competenze*) e costruiscono *statistiche, report e modelli* per la *valutazione dell'attività scientifica* di una struttura basandosi appunto sull'*output* (vale a dire le pubblicazioni scientifiche) [Galimberti, 2010, p.31].

Se adeguatamente realizzati e valorizzati, gli Archivi Istituzionali possono presentare una componente strategica per l'Anagrafe e la valutazione della ricerca a livello locale delle rispettive Istituzioni e anche a livello nazionale, svolgendo un doppio ruolo:

1) Accrescendo la *visibilità* e, quindi, l'*impatto citazionale* dei prodotti della ricerca scientifica depositati, grazie anche all'interoperabilità tecnica che viene garantita dall'utilizzo delle tecniche *Open Access*, dello standard (il modello di dati) CERIF-XML [V. sito; APPENDICE 2], dalla “Direttiva Stanca” sull'*Open Source*, dalla “Legge Stanca” sull'accessibilità⁸ e del protocollo dell'interoperabilità OAI-PMH.

⁸ “Legge Stanca” (Legge 4 del 9 gennaio 2004, pubblicata nella Gazzetta Ufficiale il 17 gennaio 2004), resa operativa col decreto attuativo DPR 1 marzo 2005 n. 75 (pubblicato nella Gazzetta Ufficiale n. 101 il 3 maggio 2005).

Il Progetto OpenAIRE (*Open Access Infrastructure Research for Europe*) [V.sito] – finanziato dalla Commissione Europea a supporto delle piattaforme di Archivi Istituzionali *Open Access* - ha fatto l'accordo con EuroCRIS⁹ [V. sito] sull'uso del modello CERIF-XML come formato standard con cui OpenAIRE raccoglie le informazioni di ricerca dai sistemi informativi per la ricerca CRIS. Il modello CERIF offre anche le funzionalità per *cross-walking* (mappatura) semantica di metadati [V. sito "EuroCRIS_Presentation"].

2) Rappresentando un'importante infrastruttura per la sperimentazione e l'applicazione congiunta delle:

- a) *nuove metriche di uso* come *Web Impact Factor*, *Usage Factor*, numero dei *download* - che possono essere applicate ai *full-text* dei contenuti depositati;
- b) *metriche citazionali*;
- c) *metriche di Social Rating* aggregate da alcune Riviste *Open Access*;
- d) più *tradizionali metriche* costruite sulla logica delle citazioni (*Impact Factor*, *H-Index*) [Baethge, 2012];
- e) *metriche per la creazione di un sistema* che sia in grado di restituire un *quadro valutativo* il più ampio e completo possibile della *performance della ricerca* di un'Università [Cassella, Bozzarelli, 2011, p.17].

Nel 2012 la Confederazione di *Open Access Repositories* (COAR) ha annunciato formalmente una cooperazione strategica con la citata associazione europea per la ricerca EuroCRIS, con l'obiettivo di promuovere l'interoperabilità tra Archivi *Open Access* e sistemi informativi di ricerca CRIS, per ottimizzare una gestione adeguata dei risultati della ricerca [V.sito: "COAR Repositories, news"].

In Italia, l'Università degli studi di Milano è stata tra le prime a integrare le funzioni dell'Anagrafe della ricerca dell'Ateneo con quelle di Archivio Istituzionale della ricerca AIR [V. sito]. In questo modo ha reso disponibili informazioni utili a verificare la rispondenza agli Standard dei contributi depositati.

L'esperienza sviluppata all'[...] AIR lavorando congiuntamente sull'incremento progressivo della auto-archiviazione dei *full.text* e su Progetti Pilota che sperimentino nuovi Indicatori e nuove Metriche e che utilizzino l'Archivio come fonte di dati per *analisi della produttività scientifica* dei ricercatori, delinea un nuovo scenario per la Valutazione, in cui alcune tradizionali debolezze (*come valutare certi risultati*, come considerare *la non linearità* di produzione in certe discipline, come *pesare il contributo individuale* nel curare opere collettive) possono trasformarsi in elementi più oggettivi, e quindi più misurabili [Galimberti (b), 2010, p.95-96].

Le Raccomandazioni del Gruppo *Open Access* della CRUI "*Open Access e la valutazione dei prodotti della ricerca scientifica*" (2009) [V.sito] evidenziano alcune considerazioni sull'importanza di collegamento di Archivi Istituzionali con le Anagrafi della ricerca (locali e nazionale) secondo determinati Standard che rendano i dati di ricerca confrontabili a livello nazionale e anche internazionale.

L'Archivio Istituzionale della ricerca di Ateneo deve essere considerato come *una componente dell'Anagrafe* [...] coinvolgendo l'Archivio aperto Istituzionale in tutte le fasi del processo di *raccolta, analisi, elaborazione, disseminazione e fruizione* dei dati relativi alla *produzione scientifica* dell'Ateneo stesso [V. sito: "*Open Access e la valutazione*", p.7].

Il collegamento tra le Anagrafi di ricerca e gli Archivi Istituzionali renderà interoperabili i dati di questi ultimi con le banche dati degli Anagrafi che devono contenere i dati su:

- a) *ricercatori*;
- b) *competenze presenti*;
- c) *risultati ottenuti in termini di pubblicazioni, libri, brevetti, partecipazione a congressi*;
- d) *attività di ricerca in corso nell'ambito di programmi e contratti*.

⁹ Un'associazione europea *not-for-profit* dedicata allo sviluppo dei sistemi informativi per la ricerca CRIS e alla loro interoperabilità.

Le succitate Raccomandazioni CRUI invitano i Nuclei di valutazione degli Atenei ad adottare uno Standard proposto dalla CRUI per il *Curriculum Vitae* [Galimberti (b), 2010, p.93], il quale può essere preso in considerazione per l'implementazione e per la gestione di Archivi Istituzionali collegati con i sistemi di valutazione della ricerca.

Tra i primi Atenei italiani collegati con i Sistemi di valutazione della ricerca si cita l'Archivio Istituzionale dell'Università degli studi di Trieste *OpenstarTs* [V.sito] che:

- a) è interoperabile con la banca dati che gestisce le Carriere dei dottorandi;
- b) è collegato con il *workflow* delle Tesi di Dottorato;
- c) è integrato con l'*iter* del conseguimento del Titolo, grazie alla collaborazione con gli Uffici ricerca e dottorati;
- d) offre uno spazio per la *conservazione* e l'*accesso a lungo termine* del materiale digitale depositato, la questione essenziale per la gestione dei dati presenti e futuri [V. sito: "KeepIt"].

I CRIS di nuova generazione sviluppati in Italia sono:

- 1) l'applicativo "*SURplus Open Archive*" [V. sito], il servizio prodotto dal Consorzio CILEA (Consorzio Interuniversitario Lombardo per l'elaborazione automatica) che fa parte dell'EuroCRIS;
- 2) il modulo "*UGOV-Ricerca*" [V. sito], il prodotto del Consorzio CINECA (formato da 56 Università italiane, 2 Enti Nazionali di Ricerca e dal Ministero dell'Università e della Ricerca, MIUR).

2.2.1.2.1. SurPlus Open Archive

Il servizio SURplus (Sistema Università Ricerca plus) *Open Archive* [De Paoli, Mornati, 2009] è un servizio integrato della suite per la *governance* dei contenuti degli Atenei ed Enti di ricerca sviluppato da CILEA, per una gestione integrata delle informazioni relative alle attività e ai prodotti della ricerca.

Tra i numerosi vantaggi che la suite SURplus può offrire ci sono:

1. l'aggiornamento dei curricula dei produttori dei contenuti che confluiscono nel sistema digitale informativo dell'Istituzione;
2. la raccolta distribuita dei dati relativi alle attività e ai prodotti della ricerca realizzando punti di raccolta centrali e unici;
3. l'interoperabilità con i sistemi informativi istituzionali ed esterni per l'integrazione di dati e funzioni;
4. la gestione integrata di tutti i dati con servizi di reportistica e monitoraggio;
5. la valutazione oggettiva dei dati basata su statistiche e indicatori di sintesi, da affiancare ai giudizi di merito;
6. la razionalizzazione e l'ottimizzazione dei processi di generazione e gestione dei dati [Mornati, 2010];
7. la conservazione delle memorie digitali, garantendo un'accessibilità globale ai dati e ai prodotti della ricerca per consentirne la loro visualizzazione e lo studio nel tempo;
8. la gestione di obiettivi, strategie e risorse, il raggiungimento dei risultati e il loro monitoraggio.

Il modulo SURplus *Open Archive* - basato sull'architettura funzionale del Software *Open Source* DSpace, il noto prodotto per Archivi Istituzionali - è sviluppato per raggiungere i seguenti obiettivi (Tabella 5):

1.	Offrire ad Atenei o Enti di ricerca i meccanismi per una flessibilità necessaria all'adattamento dei contenuti scientifici ai processi locali: la gestione integrata delle informazioni relative alle attività e ai prodotti della ricerca
----	---

	in modo da consentirne una sempre migliore quantificazione secondo metriche e indicatori condivisi [V. sito: “SURplus”].
2.	Conformare i sistemi informativi di Atenei o Enti di ricerca agli Standard internazionali che garantiscono ampia visibilità alle pubblicazioni, agli autori e alle Istituzioni accademiche e di ricerca.
3.	Realizzare un’infrastruttura di integrazione dei dati distribuiti nei sottosistemi di Ateneo, da e verso sistemi esterni, unificando, in modalità centralizzata o distribuita, le attività di inserimento dei dati di ricerca e minimizzando le possibilità di errore.
4.	Supportare il processo di creazione di Archivi digitali per la gestione delle informazioni relative alle attività di ricerca (schede progetti, contratti, brevetti, forme associative come <i>spin-off</i> ed altro) di Università ed Enti di ricerca.
5.	Realizzare l’Anagrafe dei prodotti della ricerca con funzionalità avanzate nell’ambito di CRIS (<i>Curent Reserach Information Systems</i>) a supporto della valutazione oggettiva dell’attività di ricerca e produzione scientifica basato su statistiche e indicatori di sintesi.
6.	Appoggiare la creazione e gestione degli Archivi Istituzionali dei prodotti della ricerca scientifica per mezzo del servizio SURplus <i>Open Archive</i> .
7.	Alimentare in modo automatico sezioni del portale di Ateneo relative alla ricerca e alla produzione scientifica dei docenti tramite le funzionalità informatiche <i>RSS</i> , <i>Web Service</i> o altro.

Tabella 5. Scopi del sistema SURplus (Sistema Università Ricerca plus) *Open Archive*

Tutto sommato, il modulo SURplus *Open Archive* implementa l’idea di un Archivio “che svolge la doppia funzione di Anagrafe e di Archivio aperto per la disseminazione dei prodotti della ricerca” [V. sito: “SURplus”].

Tra gli Atenei italiani, che hanno implementato i loro Archivi Istituzionali sul servizio SURplus *Open Archive*, si cita l’Archivio Istituzionale AISBERG [V.sito] dell’Università degli studi di Bergamo. Inizialmente gestito dai Servizi Bibliotecari dell’Ateneo, AISBERG adesso è integrato in una piattaforma più ampia SURplus, la cui gestione complessiva coinvolge più servizi dell’Ateneo per:

- ✓ garantire l’allineamento dei contenuti di ricerca alla struttura dati del Sito docente del Ministero;
- ✓ consentire allo stesso tempo una valutazione locale più articolata della ricerca, collegandosi coi dati provenienti da Database locali.

Questi sono i fattori principali che hanno determinato le scelte riguardanti la riorganizzazione di metadati e l’organizzazione dei contenuti dell’Archivio secondo i requisiti del servizio SURplus.

I dati di AISBERG sono inviati rispettivamente a:

- a) un modulo di *analisi statistica* e al Sito Docente del Ministero;
- b) Archivi aperti Disciplinari (*RePEc*, *ArXiv*, etc.).

La struttura dell’Archivio, ben rappresentata dall’albero delle collezioni, consente l’esatta discriminazione tra i prodotti della ricerca sottoposti a valutazione e quelli che ne sono esentati.

Anche l’Università degli Studi di Milano ha collegato i dati tratti dal proprio Archivio Istituzionale BICOCCA alla valutazione della ricerca attraverso l’uso delle funzionalità della suite SURplus, utilizzando l’Archivio Istituzionale:

- a) come Anagrafe della ricerca e
- b) anche come database bibliometrico,

fornendo tutti i dati e metadati utili alla reportistica, all’analisi statistica e alla valutazione dei prodotti della ricerca secondo varie metriche.

All’interno del modulo SURplus *Open Archive* sono previsti le seguenti funzioni (Tabella 6):

1.	il deposito (<i>submission</i>), la ricerca e il recupero in formato digitale di tutte le tipologie di lavori scientifici (anche in <i>full-text</i>): articoli su riviste, monografie e loro capitoli, contributi ad atti di convegni, brevetti, progetti, composizioni, disegni e <i>design</i> , materiale multimediale ed altre forme di produzione scientifica;
2.	l'associazione del materiale depositato con i metadati, schemi e protocolli come: <i>Dublin Core metadata</i> , XML, OAI-PHM e altri, necessari per supportare la descrizione semantica e sintattica dei documenti, e l'interoperabilità tecnica del <i>Repository</i> con altri sistemi per l'importazione ed esportazione dei dati di ricerca e la costruzione di bibliografie dei ricercatori, nonché per la predisposizione degli indicatori bibliometrici (<i>Impact Factor</i> , <i>Citation Analysis</i>) previsti dai Bandi di valutazione della ricerca;
3.	l'attuazione di aggregazioni logiche e organizzative (per Anno, per Area scientifica, per Autore, per Dipartimento, per Tipologia di pubblicazione, ecc.);
4.	l'attribuzione a ciascun lavoro scientifico di un codice identificativo persistente <i>Handle</i> ;
5.	l'estesa disseminazione e visibilità della produzione scientifica dell'Istituzione su diversi motori di ricerca (<i>Google</i> , <i>Google Scholar</i> , ecc.).

Tabella 6. Funzionalità del sistema SURplus (Sistema Università Ricerca plus) *Open Archive*

In Figura 10 vengono presentati graficamente i collegamenti tra i moduli integrati SURplus per la *governance* dei dati amministrativi: “Autenticazione”, “Gestione Anagrafica”, “*Business Intelligence*”, “*Gateway*” e quelli della ricerca gestiti entro il Modulo OA “*Open Archive*”.

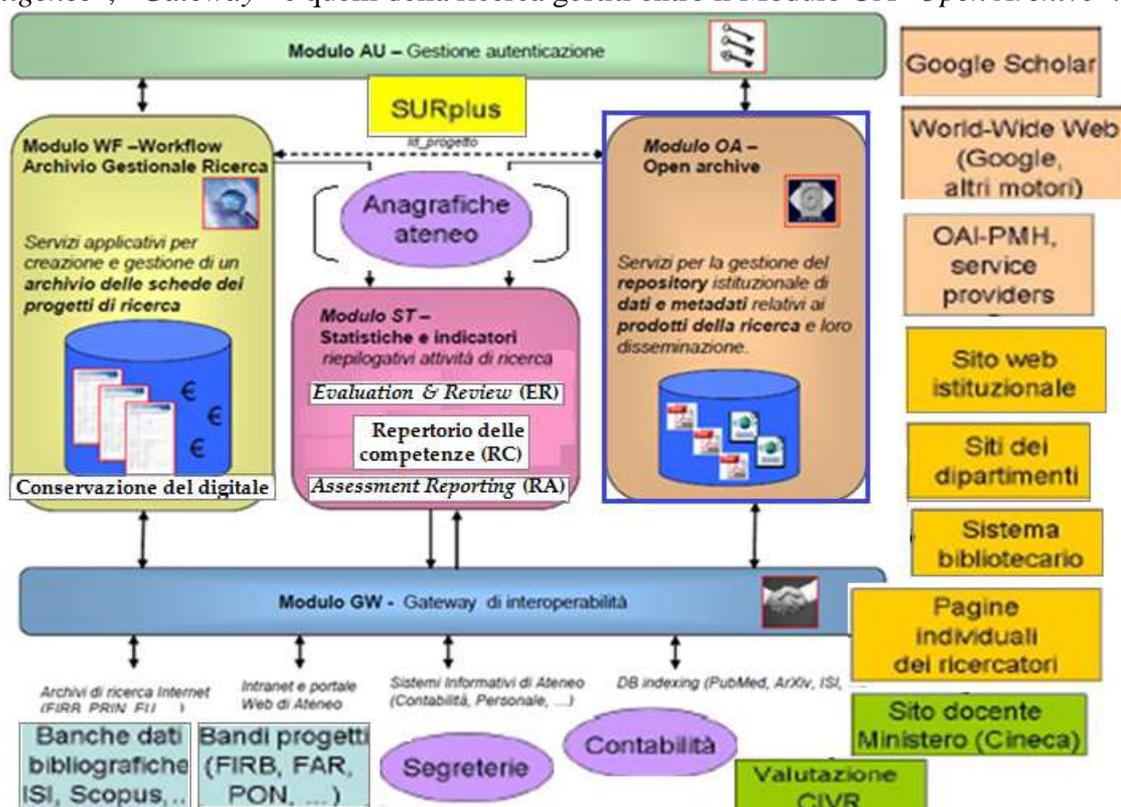


Figura 10. I collegamenti tra diversi moduli del servizio Surplus a supporto della “*governance*” dei processi e contenuti di un’Istituzione

I moduli collegati del SURplus garantiscono il supporto a un ricco apparato di ruoli e autorizzazioni, comunicazioni e accessi differenziati, per offrire una partecipazione estesa a tutte le componenti delle istituzioni/comunità che lo adottano. L’architettura del SURplus può essere

personalizzata, offrendo la possibilità di adeguare flessibilmente le funzioni e modello dei dati alle diverse esigenze istituzionali.

2.2.1.2.2. U-GOV Ricerca

La suite dei moduli integrati U-GOV - sviluppato dal CINECA (Consorzio interuniversitario per il calcolo automatico) “nasce dall'incontro tra le tecnologie più innovative, le metodologie di sviluppo più avanzate e il patrimonio di competenze e di *Best Practice* acquisite nella collaborazione con le Università italiane” [V. sito: “U-GOV”], per appoggiare la *governance* dei processi e contenuti digitali degli Atenei.

Il modulo integrato della suite - “U-GOV Ricerca” nasce come uno strumento per la creazione del “Catalogo della ricerca” e per la valutazione dei prodotti della ricerca, consentendo di attribuire una classifica di qualità ai diversi materiali di ricerca in previsione di finanziamenti ed esercizi di valutazione. In particolare, tramite la piattaforma “U-GOV Ricerca” si può:

1. creare il Catalogo della ricerca;
2. supportare la gestione integrata della produzione scientifica prodotta entro gli Atenei, rendendola operante con diverse strutture di dati;
3. valutare i prodotti della ricerca, consentendo di attribuire un *Ranking* di qualità ai diversi *output* in previsione di finanziamenti ed esercizi di valutazione [Casella, Bozzarelli, 2011];
4. qualificare la produzione scientifica come il “capitale” di un’Istituzione entro la cosiddetta “catena del valore” della “filiera della ricerca” dell’Ateneo (Figura 11), l’elemento centrale della più articolata “filiera della conoscenza” collocata tra l’alta formazione e processi d’innovazione nella filiera produttiva che ne costituiscono rispettivamente il punto di ingresso e di uscita. L’Università è l’attore protagonista dei primi due passi e interlocutrice privilegiata del mondo imprenditoriale nel trasferimento dell’innovazione verso il mercato” [V. sito: “*White Paper* - U-GOV Ricerca”, p.8].



Figura 11. La catena della “filiera della ricerca” supportata dal servizio “U-GOV- Ricerca”

La catena della “filiera della ricerca” è costituita dai seguenti moduli gestionali:

- 1) Processo di valutazione;
- 2) Disseminazione e valorizzazione dei risultati di ricerca;
- 3) Pianificazione e valutazione delle *Performance*;
- 4) Catalogazione della produzione scientifica;
- 5) Competenze, risorse, laboratori;
- 6) Progetti di ricerca.

Il modulo “U-GOV-Ricerca” raccorda

le diverse attività che portano alla creazione del valore rappresentato dai risultati della attività di ricerca in un percorso relativamente lineare nel flusso, circolare nella capacità di “re-investimento” dei risultati stessi [V. sito: “White Paper - U-GOV Ricerca”, p.9].

“UGOV-Ricerca” può interagire con i seguenti moduli funzionali della suite U-GOV dedicati alla gestione dei principali processi amministrativi interni all'Ateneo (Figura 12):

1. UGOV- Pianificazione e controllo;
2. UGOV- Contabilità;
3. UGOV- Didattica e studenti;
4. UGOV- Risorse umane;
5. UGOV- Gestione documentale.

L'uso di Standard aperti, in particolare lo Standard JEE nello sviluppo delle componenti applicative, consente di potenziare la comunicazione fra i diversi moduli ed offrire applicazioni distribuite, sicure e interoperabili [V. sito: “U-GOV, Tecnologia e Architettura”].

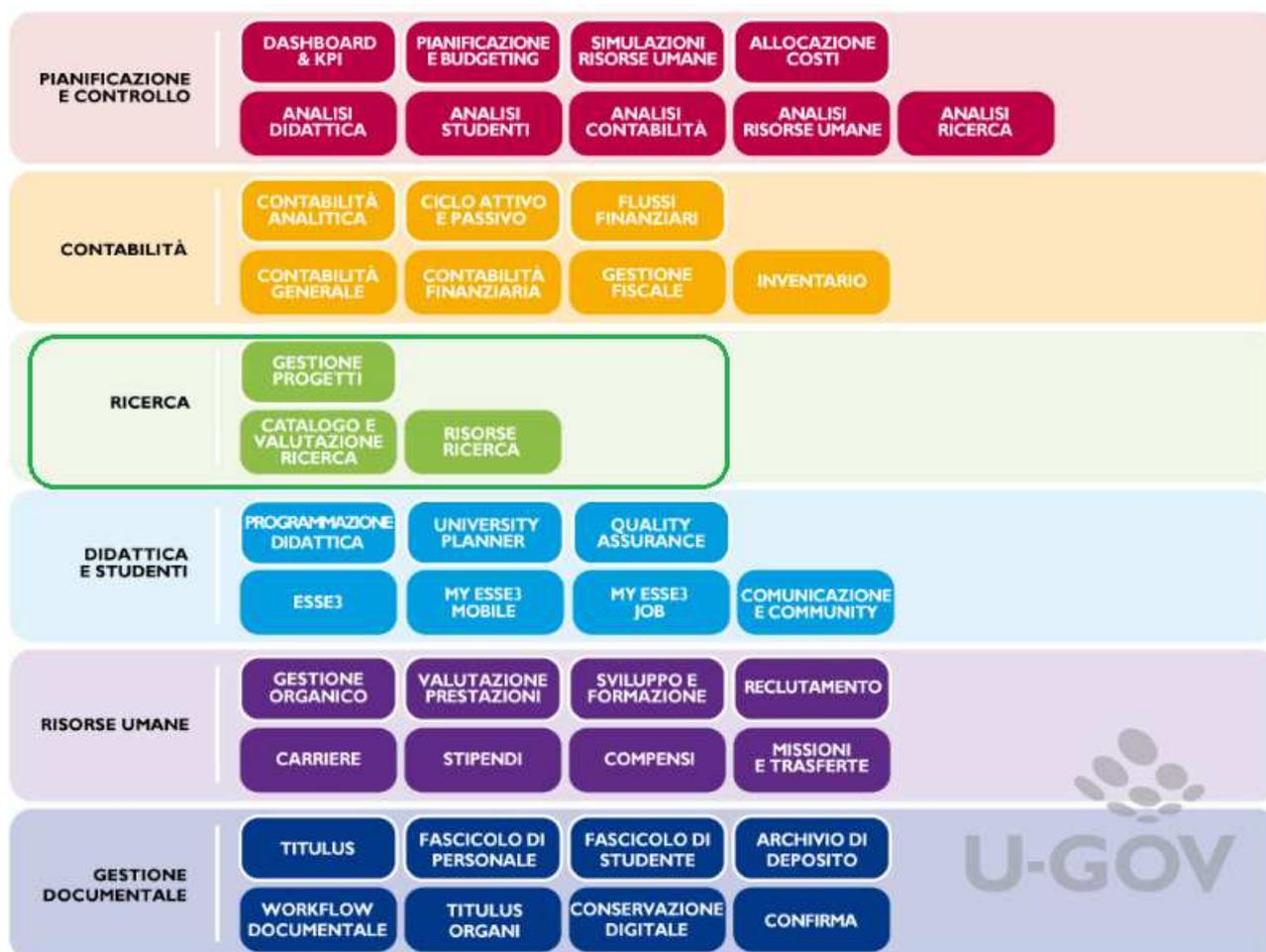


Figura 12. L'architettura del sistema informatico integrato U-GOV
[V. siti: "U-GOV"; "U-GOV per la governance degli Atenei"]

Come presentato in Figura 11, l'infrastruttura tecnologica della suite U-GOV consente di attivare vari livelli di integrazione tra le diverse aree funzionali d'Ateneo ("Pianificazione e Controllo", "Contabilità", "Ricerca", "Didattica e Studenti", "Risorse Umane", "Gestione Documentale") utilizzando il potenziale fornito dalla propria architettura.

Il modulo "U-GOV-Ricerca" è articolato in seguenti sottomoduli:

1. "Catalogo e Valutazione dei Prodotti della Ricerca", che consente di:
 - a) gestire le pubblicazioni dei ricercatori dell'Ateneo trasferendole direttamente a Portale del Ministero;
 - b) attribuire un *Ranking* di qualità ai prodotti in previsione di finanziamenti e valutazioni nazionali.
2. "Risorse della Ricerca". Tale sottomodulo consente di:
 - a) gestire e censire (verificare il corretto impiego delle risorse investite) tutte le entità coinvolte nella Ricerca dell'Ateneo;
 - b) fare analisi sui dati (compreso il *know-how* acquisito);
 - c) pianificarne i temi;
 - d) creare una "vetrina" per agevolare la collaborazione con Industrie e Aziende private.
3. "Gestione dei Progetti di Ricerca". Tale sottomodulo consente di:
 - a) censire e gestire i dati scientifici dei Progetti di Ricerca (sia dal punto di vista operativo che finanziario);
 - b) collegare i dati scientifici ad altri moduli della suite U-GOV [Luglio, 2011].

Al momento (2013) in Italia 36 Università utilizzano il Catalogo "UGOV-Ricerca" come lo strumento per l'inserimento della produzione scientifica da parte dei propri ricercatori. L'Archivio Istituzionale dell'Università di Torino "Publication Open Repository Torino" (PORTO) [V. sito] - sviluppato sul *Software Open Source EPrints* nel 2010 - è stato la naturale evoluzione del progetto "Archivio prodotti della Ricerca" basato sul modulo "UGOV-Catalogo della Ricerca", a cui sono stati affiancati i meccanismi per la comunicazione con la piattaforma dell'Archivio PORTO (Figura 13).

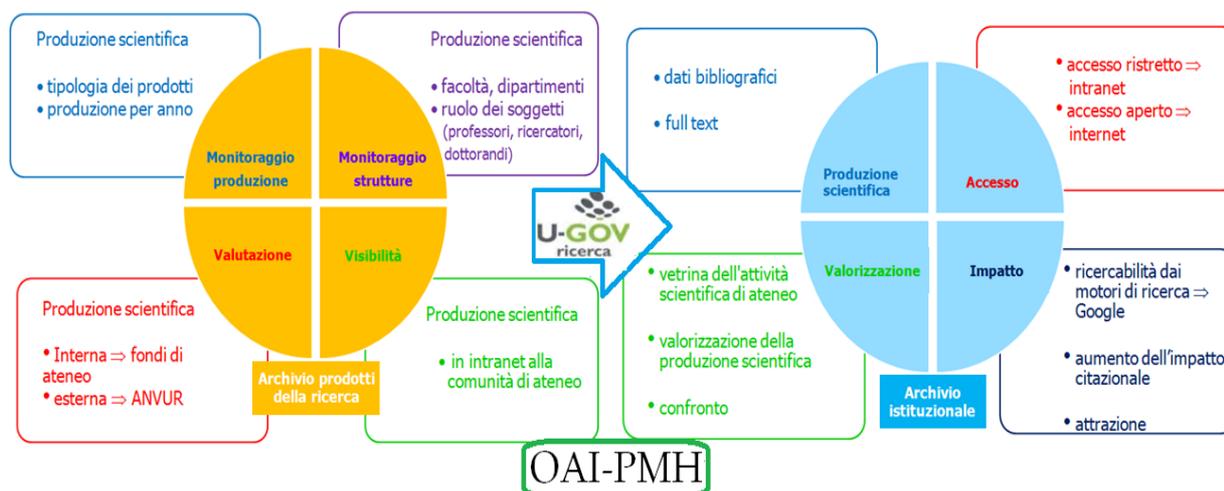


Figura 13. La comunicazione tra i moduli "UGOV-Catalogo della Ricerca" e l'Archivio Istituzionale del Politecnico di Torino

I dati sulle pubblicazioni vengono inseriti una volta sola dai ricercatori dell'Ateneo e poi esportati da U-GOV - tramite il protocollo OAI-PMH e un applicativo sviluppato dal Politecnico di Torino - che carica i metadati e i documenti associati in PORTO, nonché i dati relativi alle gestione del *Copyright* che non sono ancora ammessi in U-GOV. Il gestionale PORTO contiene, quindi, una copia esatta delle pubblicazioni presenti in U-GOV ed è strutturato per:

- offrire un'interfaccia pubblica di U-GOV per la consultazione in rete di tutti i prodotti della ricerca dell'Ateneo presenti in U-GOV;
- indicizzare in modo raffinato i contenuti di U-GOV, effettuando raggruppamenti in categorie dei documenti;
- partecipare ai processi di valutazione dei prodotti della ricerca ai fini di un loro impiego corretto e ottimizzato nella complessa catena del valore scientifico dell'Ateneo;
- raccogliere le statistiche sui download dei "pdf";
- controllare il *Copyright* grazie al supporto di uno *staff* dedicato alla valutazione dei diritti di archiviazione ed accesso di ciò che viene reso disponibile in *Open Access*.

Tramite "UGOV-Catalogo della Ricerca" i ricercatori possono utilizzare i dati inseriti:

- sia per finalità Ministeriali (sul sito individuale del docente "*sitouniversitario.cineca.it*", ma anche su sistemi dedicati all'esercizio di valutazione della ricerca nazionale) che
- interne all'Ateneo (relativamente al portale di Ateneo, Archivi Istituzionali, pagine di Dipartimenti).

"UGOV-Catalogo della ricerca" e PORTO *Eprints* lavorano come due gestionali separati, anche se tra loro collegati e interoperabili. Lo scopo di creare un'interfaccia pubblica per "U-GOV" tramite l'Archivio Istituzionale è quello di utilizzare un unico gestionale. Questa soluzione permette anche di ottimizzare gli sforzi di sviluppo e di manutenzione dei due applicativi.

2.3. Etica nel rispetto di contratti editoriali e proprietà intellettuale

Gli aspetti legati alla proprietà intellettuale sono di primaria importanza da considerare durante l'auto-archiviazione dei lavori in un Archivio Istituzionale, in quanto gli autori devono essere convinti del valore dei propri lavori al punto da poter controllare quali utilizzi ne vengono fatti. Nella maggior parte dei casi, pubblicando i propri lavori di ricerca sulle Riviste tradizionali, gli autori cedono il controllo esclusivo dei medesimi agli editori. Ciò causa una serie di limiti che riguardano:

- ✓ la diffusione dell'informazione scientifica verso un grande pubblico dei lettori;
- ✓ l'impatto della medesima.

La mancanza di un accesso diffuso alla ricerca scientifica delinea infatti un freno per la ricerca stessa, e può comportare:

- ✓ una ripetizione degli sforzi nella produzione delle stesse idee di ricerca;
- ✓ un'inutile dissipazione di tempo, di energie e di fondi pubblici.

Se l'autore può essere posto al centro del circuito della comunicazione scientifica, tre sono le tipologie di rapporti legati alla proprietà intellettuale:

- 1) Autore-Editore,
- 2) Autore-Istituzione,
- 3) Autore-Lettori.

Secondo la banca dati “*Sherpa/RoMEO*” (*Publisher copyright policies & self-archiving*) [V.sito: “*Sherpa_Politiche*”] - che indica quali sono le Politiche di una serie di Editori commerciali rispetto all'auto-archiviazione dei lavori di ricerca negli Archivi aperti da parte degli Autori, risulta che molti grandi editori (69%) permettono l'auto-archiviazione di file “.pdf”, nella versione finale da loro prodotta - in un Archivio aperto, nell'accezione della strategia “*Green Road*” dell'*Open Access*, o sul sito personale dell'Autore.

Quindi, nell'Archivio aperto *Open Access*, accanto al materiale pubblicato nelle Riviste *Open Access*, possono essere depositati anche i *full-text* dei lavori pubblicati sulle Riviste tradizionali, che permettono l'auto-archiviazione.

Nel 2012 la *Confederation of Open Access Repositories (COAR)* ha annunciato la formazione di una *Task Force* [V.sito: “*Confederation of Open Access Repositories_Task Force*”] dedicata alle Licenze e agli Accordi con gli editori a favore dell'Accesso Aperto, proponendosi di rivedere e valutare gli accordi in atto tra editori ed enti di ricerca che prevedano Clausole sull'*Open Access*. Considerando, quindi, che le Politiche editoriali possono cambiarsi nel tempo [V. sito: “*Blog di R. Poynder*“], è importante che l'Archivio Istituzionale sia affiancato dal personale che si occupi della gestione delle problematiche relative al *Copyright* in relazione alle opere dei depositi e ne tenga traccia delle modifiche, per poter modificare le *Policy* del deposito che devono affinare gli Archivi Istituzionali.

Quanto alle Riviste “*Open Access*”, i loro Editori non richiedono agli autori la cessione esclusiva dei diritti/del *Copyright*, proponendo un contratto sotto una *Licenza non esclusiva* che permette all'autore di *mantenere alcuni* dei *diritti* di sfruttamento economico fra cui quello di auto-archiviare il proprio lavoro in un Archivio aperto ad Accesso Aperto. Per gli autori che auto-archiviano i loro lavori nei medesimi Archivi, sarebbe utile specificare quali sono i diritti che essi concedono ai lettori e a quali condizioni. L'autore che non ha ceduto tutti i diritti ad un editore può decidere se *elargire* ai lettori *tutti i diritti* di utilizzazione economica o soltanto *alcuni*, tramite le Licenze *Open Content* (es. le Licenze *Creative Commons/CC* [V.sito], *GNU General Public License/GPL*) [De Robbio, 2012].

Le licenze CC sono dei *contratti veri e propri* mediante i quali il titolare dei diritti concede una serie più o meno ampia di *diritti ai fruitori dell'opera*. In contrasto con il Diritto d'autore tradizionale che si basa sulla formula Tutti i diritti riservati, CC si basa su una gestione più

flessibile del Diritto d'autore, quella fondata sulla formula *Alcuni diritti riservati*. Principio ispiratore è senza dubbio quello delle licenze GPL, tuttavia le licenze CC non nascono per proteggere il Software ma altri tipi di opere dell'ingegno: musica, film, fotografie, testi ecc. [V. sito: "Diritto d'autore"].

Le licenze *Creative Commons* forniscono agli autori dei mezzi affinché possano specificare facilmente delle autorizzazioni o dei permessi agevolmente leggibili e comprensibili dagli umani e/o dalle macchine.

Per quanto riguarda gli Editori italiani, l'articolo 42 della Legge 22 aprile 1941, n. 633 sul "*Diritto d'autore*" stabilisce che l'autore di un articolo su rivista ha diritto a riprodurlo altrove, purché citi gli estremi della prima pubblicazione, a meno che non sia stato esplicitamente stabilito il contrario. Gli editori italiani non si sono ancora pronunciati rispetto all'auto-archiviazione negli Archivi Istituzionali.

Per essere del tutto certi di poter pubblicare un articolo in un Archivio aperto *Open Access*, vale comunque sempre la pena di rileggere il contratto o di chiedere direttamente all'editore. In assenza di clausole esplicite nel contratto o in assenza di contratto, l'autore può auto-archiviare il suo articolo in un Archivio aperto *Open Access*.

Nel caso in cui l'Editore rifiuti di concedere all'autore la Licenza di depositare il suo articolo in modalità ad *Accesso Aperto*, le soluzioni maggiormente adottate per aggirare questo problema sono le seguenti:

1) l'auto-archiviazione dell'articolo nella sua forma di "*pre-print*", ossia nello stato dell'articolo *precedente la revisione da parte del comitato editoriale*. Dopo la pubblicazione dell'articolo si può depositare separatamente il file di corrigenda, connesso al "*pre-print*". Intanto è sempre opportuno chiedere all'eventuale futuro editore se l'auto-archiviazione può pregiudicare la pubblicazione dell'articolo.

2) Se l'editore impedisse esplicitamente anche l'auto-archiviazione del "*pre-print*", l'autore potrebbe comunque depositarlo applicando - però - delle *restrizioni* all'accesso.

La *cessione in forma esclusiva dei Diritti* preclude una serie di riutilizzi delle proprie opere (sia a scopo didattico che scientifico), per cui deve essere richiesta l'autorizzazione agli Editori i quali possono concederla o meno (a titolo oneroso o meno). L'autorizzazione per riutilizzi didattici o scientifici (a scopi non commerciali) da parte dell'Autore delle proprie opere può essere richiesta tramite la cosiddetta "Addenda" o l'ADDENDUM [V.sito] da includere nel contratto dell'Editore;

3) "*embargo*": l'auto-archiviazione del *full-text* dell'articolo è consentita a partire da un determinato tempo, successivo alla pubblicazione sulla Rivista. Intanto viene depositata la descrizione dell'articolo attraverso i metadati.

Una ricca guida informativa per conoscere il quadro legale (*Licensing Terms, Access Control, Ownership rights, Exploitation rights: Reproduction right; Communication right, Distribution right, Transformation right*) entro il quale possono compiersi diritti morali e patrimoniali e Politiche di *Copyright* in relazione al deposito dei lavori di ricerca negli Archivi Istituzionali è offerta dal Servizio online "*Padova Digital University Archive*" – "Politiche di *Copyright*" [V.sito].

Il quadro legale determina e regola i diritti individuali e questi ultimi regolano i permessi concessi dai titolari ad altri; infine, i permessi garantiti dai titolari possono determinare, caso per caso, il consenso al compimento di singole azioni da parte dell'utilizzatore (*user*).

Se l'autore ha già sottoscritto (o prevede di sottoscrivere) con un Editore un contratto di cessione dei propri diritti o ha trasferito all'Editore il *Copyright*, prima di procedere all'auto-archiviazione:

1. per i *Contratti già firmati che non prevedono il consenso all'auto-archiviazione*, l'autore dovrebbe richiedere all'Editore che il Contratto venga modificato in modo tale da consentire l'auto-archiviazione, inserendo questo Modello di clausola:

*Delego a [l'editore o la rivista] tutti i Diritti di vendere o di cedere il testo (su carta o online) del mio articolo [titolo dell'articolo].
Conservo solo il diritto di diffonderlo gratuitamente a fini scientifici o didattici, in particolare il diritto **di auto-archiviarlo pubblicamente online sul Web***

2. Per i Contratti che devono ancora essere sottoscritti, l'autore dovrebbe inserire una clausola che preveda il consenso all'auto-archiviazione.

È di vitale importanza per chi gestisce l'Archivio Istituzionale abbia una buona conoscenza delle tematiche legate al Diritto d'Autore. Tutte le informazioni utili a supporto dell'Autore e del Lettore devono essere approntate negli appositi FAQ dell'Archivio Istituzionale e anche nelle Guide informative [V. Linee guida per l'Autore] all'auto-archiviazione per spiegare agli autori come comportarsi nel caso dell'auto-archiviazione di un certo materiale di ricerca (Figura 14).



Figura 14. Le informazioni sul “Diritto d’Autore” sul sito dell’Archivio della Ricerca AIR dell’Università degli Studi di Milano [V.sito].

2.4. Conservazione a lungo termine dei contenuti digitali

Le “Linee Guida per gli Archivi Istituzionali” (CRUI, 2012) [V.sito] accentuano l'importanza della conservazione (preservazione) permanente della conoscenza scientifica a futura memoria, che è un tema strategico per la comunicazione dell'informazione scientifica nel Ventunesimo secolo, l'era dei contenuti digitali. La conservazione del digitale (*digital preservation*) può essere raggiunta tramite l'applicazione dello Standard OAIS (*Open Archival Information System*), lo Standard più diffuso e usato per definire il modello delle funzioni coinvolte nella conservazione digitale a lungo termine, aggiornato e approvato formalmente come lo Standard ISO 14721:2012 [V.sito]. Il modello OAIS ha due obiettivi principali:

- a) stabilire una terminologia standard per descrivere le caratteristiche degli Archivi che assolvono la funzione di archiviazione;
- b) stabilire un livello minimo di funzionalità per l'archiviazione negli Archivi.

La conformità di un qualsivoglia Archivio digitale allo standard OAIS, al quadro normativo nazionale, alle strategie e best practices di conservazione a lungo termine rappresenta un indice rilevante della sua qualità.

Gli strumenti a supporto delle Politiche di preservazione sono i seguenti:

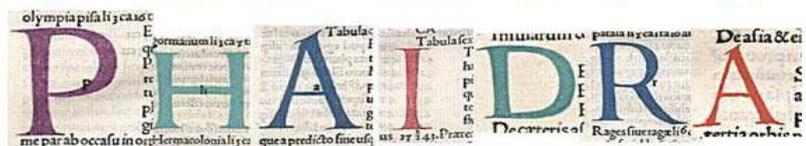
- a) la definizione di formati di file da depositare - compatibili con l'archiviazione a lungo termine e le garanzie offerte per assicurare l'integrità dei materiali ospitati;
- b) l'utilizzo delle tecniche di gestione formati (conversione, migrazione, emulazione);
- c) attribuzione ai contenuti depositati degli identificatori basati sui link permanenti;
- d) l'uso dei sistemi di *storage*, *backup*, *business continuity*, *disaster recovery*, reportistica (*log*);
- e) la gestione metadati per la conservazione e *refresh* dei dati;
- f) la gestione metodologica all'accesso ai dati (sicurezza, autenticazione, autorizzazione, sorveglianza);
- g) l'applicazione della gestione dell'autenticità dei dati (integrità, firma digitale).

“I costi legati alla preservazione permanente sono estremamente elevati. Si raccomanda dunque di selezionare accuratamente il materiale per cui si vuole garantire la persistenza” [V. sito: “Linee Guida per gli Archivi Istituzionali”, p.28]. Inoltre, va considerato che il tema della “conservazione del digitale” è un tema complesso dalle molteplici implicazioni che devono basarsi su sinergie: tecniche, economiche, organizzative e giuridiche.

Nel contesto degli Archivi Istituzionali la realizzazione di efficaci e stabili azioni di conservazione dei documenti digitali a lungo termine è un chiaro segno di successo dei medesimi Archivi. Ad esempio, gli Archivi Istituzionali dell'Università di Padova e l'Università di Vienna - sviluppati sulla piattaforma PHAIDRA (*Permanent Hosting, Archiving and Indexing of Digital Resources and Assets*) [V.sito] (Figura 15), la personalizzazione del Software *Open Source* Fedora - hanno attuato diverse azioni a supporto della conservazione dei materiali digitali a lungo termine [Höckner, Budroni, 2011].



Phaidra - Permanent Hosting, Archiving and Indexing of Digital Resources and Assets at the University of Vienna



Che cos'è PHAIDRA

Acronimo di *Permanent hosting, archiving and indexing of digital resources and assets*, PHAIDRA è un sistema di gestione delle collezioni digitali con funzioni di archiviazione a lungo termine.

Facile! La descrizione e l'indicizzazione degli oggetti digitali sono semplici e veloci. La citabilità di ciascuna risorsa caricata è garantita da un link permanente.

Affidabile! La piattaforma è basata sul software libero *Fedora Commons*. I formati adottati per i metadati e gli oggetti digitali garantiscono l'archiviazione a lungo termine e l'interoperabilità con altri sistemi.

Sicuro! Grazie alla definizione delle condizioni d'uso e di un sistema di licenze per gli oggetti caricati.

Formati di file compatibili con l'archiviazione a lungo termine in Phaidra

Testo		Audio	
.pdf (PDF/A), .tei	consigliati	.wav	consigliato
.tex, .html	possibili	.mp3	possibile
.doc, .ppt, .indd, .qxd	sconsigliati	.wma, .aac, .ogg	sconsigliati
Immagine		Video	
.tiff	consigliato	.avi, mpeg2	consigliati
.jpg	possibile		possibile
.psd, .bmp, .gif, .png	sconsigliati	.wmv, .flv, .mov	sconsigliati

Figura 15. La piattaforma PHAIDRA: un sistema di gestione delle collezioni digitali con le funzioni di *archiviazione a lungo termine*

L'architettura tecnica di PHAIDRA:

- a) è costruita sul Software *Open Source* Fedora che viene utilizzato per la creazione degli Archivi aperti;
- b) garantisce la citabilità di ogni risorsa tramite un link permanente;
- c) prevede il caricamento dei formati dei file compatibili con l'archiviazione a lungo termine;
- d) garantisce l'archiviazione a lungo termine e l'interoperabilità con altri sistemi.

A partire dalla metà degli anni Duemila, ovvero non appena gli Archivi Istituzionali hanno superato la fase sperimentale dei primi anni, diversi Progetti, finanziati dall'Unione Europea, hanno mirato a realizzare un'architettura per la conservazione degli oggetti digitali a lungo termine. I più importanti Progetti in questa direzione, che hanno predisposto una documentazione specifica sulle Politiche di conservazione a lungo termine e la loro pianificazione negli Archivi aperti, sono:

- 1) 2013-2015: ODIN (ORCID and DataCite Interoperability Network) [V. sito];
- 2) 2013 – 2015:PRELIDA Project – Preserving Linked Data [V. sito];
- 3) 2012-2015: Benchmark DP Project – Systematic approach to assess and compare digital preservation processes, systems, and organizational capabilities Systematic approach to assess and compare digital preservation processes, systems, and organizational capabilities [V. sito];
- 4) 2011- 2015: TIMBUS - Digital Preservation for Timeless Business Processes and Services [V. sito];
- 5) 2011- 2015: APARSEN (Alliance for Permanent Access to the Records of Science Network of Excellence) *Project* [V. sito].
- 6) 2011-2014: WF4Ever Project - Preservation of Scientific Workflows [V. sito];
- 7) 2011 – 2014: SCIDIP-ES - SCIENCE Data Infrastructure for PRESERVATION Project (2011 - 2014) [V. sito];
- 8) 2009 – 2010: *KeepIt Project*, Setting institutional repositories on the path to digital preservation [V. sito];
- 9) 2007 – 2011: SHAMAN - Sustaining Heritage Access through Multivalent Archiving: [V. sito];
- 10) 2006 – 2009: *Preservation and Long-term Access through NETWORKED Services* (PLANETS) [V. sito].

I Progetti menzionati hanno contribuito a elevare il livello di consapevolezza dell'importanza della conservazione digitale tra i vari *stakeholders* della comunicazione dell'informazione scientifica e tra i *Repository Managers*. Tuttavia, le pratiche di *Digital Preservation* in molti Archivi Istituzionali restano ancora confinate ad operazioni di semplice manutenzione, in parte perché spesso è problematico sostenere i costi elevati delle attività di conservazione del digitale, in parte per la mancanza di sinergie che circoscrivano i ruoli e le funzioni dei vari attori nelle Politiche di conservazione.

2.5. Data Management (and Process) Plan per la programmazione della gestione di contenuti

Ogni gestore dell'Archivio Istituzionale è consapevole che la corretta pianificazione e gestione dei contenuti digitali del medesimo Archivio ricopre un ruolo centrale per il successo di quest'ultimo in una prospettiva di medio-lungo termine. In questo contesto, la stesura di un *Data Management and Process Plan* avrà lo scopo di formalizzare la programmazione della gestione di:

- 1) tutto il ciclo di vita di contenuti digitali (es. articoli, set di dati, immagini, flussi di dati, oggetti multimediali);

Il ciclo di vita dell'informazione comprende la *creazione*, la *trasmissione*, il *processo*, la *gestione* e l'*uso* dell'informazione stessa attraverso la *condivisione*, la *modificazione*, l'*indicizzazione* ai fini del suo recupero, la *conservazione* e l'*immagazzinamento*, ma anche e soprattutto l'*apprendimento* e l'*istruzione* [Floridi, 2010, pp. 85, 86];

2) processi che coinvolgono diversi agenti (es. autori, istituzioni, personale tecnico-amministrativo), servizi, tecniche, località geografiche e altre entità - da cui dipende la qualità dell'intera filiera di gestione e monitoraggio dei contenuti dell'Archivio (dalla raccolta fino alla conservazione e il riuso dei contenuti) (Figura 16).

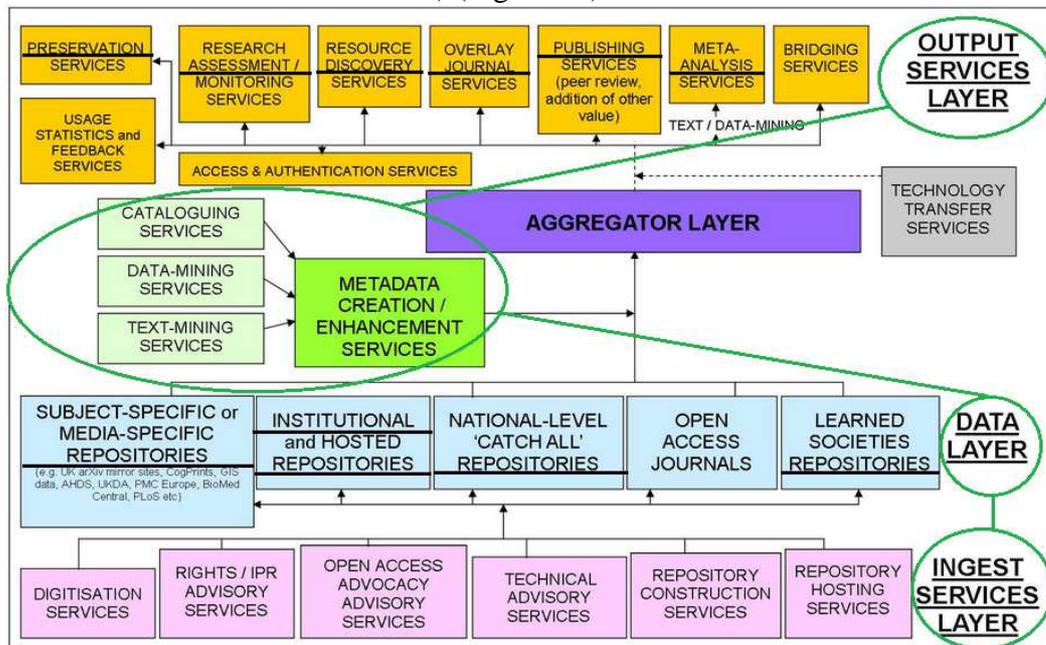


Figura 16. La filiera del ciclo di vita dei contenuti di un Archivio e i processi coinvolti [V. sito: “Enabling Open Scholarship (EOS)”¹⁰]

Secondo la definizione del servizio “Digital Curation Center” [V.sito], *Data Management Plans (DMPs)*

typically state what data will be created and how, and outline the Plans for sharing and preservation, noting what is appropriate, given the nature of the data and any restrictions that may need to be applied [V.sito “Digital Curation Center”].

Oltre a supportare la gestione del ciclo di vita dei contenuti digitali di un Archivio, lo sviluppo di un DMP può essere indicato anche per:

- a) partecipare ad un Progetto di ricerca;
- b) le parti che riutilizzano i contenuti di ricerca.

Il “Digital Curation Center” offre i seguenti strumenti (metodi tecnici) e le linee guida per lo sviluppo dei contenuti di un DMP, quali sono:

1. “DMP Online”: uno strumento web flessibile che guida la creazione di DMPs personalizzati in base ai diversi contesti;
2. “Funders’ Data Plan requirements”: la sintesi delle aspettative dei finanziatori di ricerca per la gestione dei dati e la condivisione dei DMPs;
3. “Checklist for a Data Management Plan”: un elenco completo delle informazioni che i ricercatori sono invitati a includere nei DMPs;
4. “DMP checklist leaflet”: una sintesi della Checklist;

¹⁰ *EnablingOpenScholarship (EOS)* è un'organizzazione a supporto di Università e istituti di ricerca in tutto il mondo. Il servizio online EOS, supportato dall'Università di Liège (Belgio), rappresenta sia un luogo di informazioni che un forum per sollevare e discutere le questioni intorno alla missione di Università e Istituti di ricerca moderni, in particolare per quanto riguarda la creazione, diffusione e conservazione dei risultati della ricerca.

5. “*How to Develop a Data Management and Sharing Plan*”: una guida che illustra le aspettative dei finanziatori dei Progetti di ricerca e le informazioni su come poter soddisfare le medesime;

6. “*Guidance and examples*”: consigli pratici per scrivere un DMP.

La formulazione di un DMP deve basarsi sulla *Policy* dell’Archivio Istituzionale preventivamente ben definita, che comprende la descrizione di diversi processi a supporto dei contenuti digitali.

Tra i punti principali di un DMP c’è la definizione dei seguenti processi:

- a) raccolta dei dati ovvero del deposito dei contenuti (documenti, dati, materiale multimediale) in rispetto all’ontologia degli oggetti digitali previsti per il deposito nell’Archivio;
- b) creazione di metadati;
- c) ritrovamento, identificazione e selezione dei contenuti;
- d) visualizzazione dei contenuti;
- e) preservazione dell’integrità (completezza) dei contenuti digitali, compresa l’informazione sulla loro provenienza, la definizione dei formati e delle Politiche di conservazione;
- f) condivisione dei contenuti con altri sistemi, e il loro riuso dai terzi.

Attraverso la stesura di un *Data Management and Process Plan* devono essere quindi formalizzati gli strumenti a supporto della qualità, sicurezza, persistenza a lungo termine e del riuso corretto delle informazioni digitali dell’Archivio Istituzionale. La negligenza verso una sistematica programmazione dei contenuti digitali e i processi a loro supporto può provocare decisioni gestionali errate e, come effetto, portare ad una cattiva gestione, perdite di dati o allo spreco di altre risorse coinvolte nella gestione dei contenuti dell’Archivio.

2.6.Trusted Digital Repositories: verso la sostenibilità e l’affidabilità delle infrastrutture di ricerca

Il *Research Library Group* (RLG) - tramite il Rapporto “*Trusted Digital Repositories: Attributes and Responsibilities*” (2002) [V. sito] - è stato primo a definire gli *attributi e le responsabilità* che denotano la prassi di un “*Trusted Digital Repository*” (TDR), un Archivio digitale che offre l’*accesso affidabile a lungo termine* alle proprie risorse digitali. I sette attributi di un TDR sono:

- 1) la conformità al modello di riferimento per la conservazione a lungo termine *Archival Open Information System* (OAIS);
- 2) la responsabilità amministrativa;
- 3) la sostenibilità organizzativa;
- 4) la sostenibilità finanziaria e tecnologica;
- 5) l’adeguatezza procedurale;
- 6) la sicurezza del sistema;
- 7) la responsabilità di procedure dell’Archivio (*Repository*),

tutto ciò in riferimento alla tematica della preservazione dei contenuti digitali a lungo termine.

Il succitato Rapporto include anche le Raccomandazioni per lo sviluppo di un sistema collaborativo di Archivi digitali, attraverso cui possono essere condivisi non solo materiali di ricerca ma anche le informazioni sulla preservazione digitale a lungo termine e gli aspetti che riguardano i Diritti di proprietà intellettuale. TDR è un presupposto importante per la certificazione di un Archivio; va sviluppato conformandosi ai passi indicati nel documento “*Trustworthy Repositories Audit & Certification: Criteria and Checklist*” TRAC [V. sito] e, di recente, nello standard ISO 16363:2012 “*Space data and information transfer systems -- Audit and Certification of Trustworthy Digital Repositories*” (TDR) [V. sito].

Diverse tematiche collegate a TDR sono state approfondite durante la Conferenza Internazionale “*Trusted Digital Repositories & Trusted Professionals*” [V. sito] tenutasi nel 2012 a Firenze. Le tematiche della Conferenza hanno ben curato gli aspetti relativi a:

- a) la preservazione del digitale negli Archivi di ricerca;
- b) il modello standard per la conservazione a lungo termine OAIS, in connessione alla gestione integrata dei diritti di accesso e alla certificazione di TDR;
- c) le linee guida “*Data Seal of Approval (DSA)*” per l'applicazione e la verifica degli aspetti di qualità per quanto riguarda la creazione, l'archiviazione e il (ri)utilizzo di dati di ricerca digitali;
- d) il DIN Standard 31644 "Criteri per Archivi digitali affidabili" e la certificazione “*Nestor*” (la rete di competenza tedesca per la conservazione digitale) per supportare l'attestazione di qualità di TDR;
- e) il “*European Framework for Audit and Certification*” degli Archivi digitali;
- f) la Legislazione Nazionale sulla conservazione digitale e alle buone pratiche per la costruzione di Archivi di qualità sostenibili nel tempo.

Tra gli strumenti per la certificazione della qualità degli Archivi digitali è da citare anche la certificazione DINI [V. sito]. Per richiedere online tale certificazione, un Archivio deve conformarsi ad una serie di requisiti minimi. Una volta ottenuta la certificazione, il sigillo DINI può far parte della pagina web dell'Archivio garantendo così ai suoi utenti l'affidabilità e qualità del Servizio (Figura 17).

Figura 17. Archivio Istituzionale “*EconStor*” [V.sito] dell'Università Humboldt (Berlino) con la certificazione DINI

L'obiettivo dell'Archivio certificato, cioè dell'Archivio di qualità approvata, è quello di offrire agli utenti un sistema sicuro per il deposito, la conservazione e la ricerca dei contenuti, secondo opportuni piani e criteri di successo formalizzati e attuati tramite le migliori prestazioni dell'Archivio come servizio.

2.7. Conclusioni

Gli Archivi aperti Istituzionali sono una realtà da oltre un decennio, e risultano molto utilizzati per gli scopi di auto-archiviazione e di ricerca dei materiali scientifici in diverse discipline. Per confermare tale fatto, è doveroso citare una recente indagine internazionale a cui hanno partecipato circa 1700 ricercatori (per la maggior parte fisici) [Nicholas et al., 2012]. La medesima indagine ha rivelato che gli Archivi Istituzionali si sono ricavati uno spazio di rilievo importante nelle comunità di riferimento e sono molto utilizzati, specialmente dai giovani ricercatori.

Per raggiungere i livelli di qualità, la programmazione e l'implementazione di Archivi Istituzionali ci si deve conformare ad una serie di Linee Guida, Raccomandazioni e Standard promossi e ampiamente riconosciuti a livello nazionale e internazionale.

Gli Archivi appoggiati sui servizi informatici “*SURplus Open Archive*” e “*U-GOV-Ricerca*” hanno la possibilità di basare i loro contenuti su meccanismi che facilitano la *data governance* ovvero il governo dei contenuti digitali creati entro Atenei o Enti di ricerca. Tali meccanismi coinvolgono l'insieme di metodologie e attività necessarie per standardizzare i dati e il loro significato rendendoli consistenti tra loro ed integrandoli non solo nei diversi sistemi dell'Ateneo ma anche in quelli esterni. La *governance* dei contenuti digitali non si riferisce solo ad una soluzione tecnologica, al contrario; la *data governance* è una combinazione di persone, processi e tecnologie che garantiscono l'alta qualità dei dati, e quindi fanno sì che le informazioni ad alto valore, molto accurate, possano essere la base su cui costruire l'efficienza di un sistema [V. sito: “*Data Manager Online*”].

Affinchè un Archivio possa raggiungere i livelli di qualità, esso deve conformarsi ai requisiti di:

- 1) etica, nel rispetto di contratti editoriali e proprietà intellettuale relativamente ai contenuti da depositare;
- 2) conservazione a lungo termine dei contenuti depositati.

Per di più, un Archivio Istituzionale deve assicurare:

- a) la comprensione qualitativa della/e propria/e *Policy*;
- b) la diffusione della conoscenza sulla/e medesima/e *Policy* tra le Istituzioni e utenti;
- c) l'attuazione dei requisiti a supporto della qualità dei contenuti digitali, quali l'affidabilità (*reliability*), la sicurezza (*safety*) e la gestione dei rischi (*risk management*) collegati ai diversi processi del sistema Archivio.

Per attuare meccanismi e processi a supporto del ciclo di vita dei contenuti digitali dichiarati nella(e) *Policy* dell'Archivio Istituzionale, è opportuno che, a partire dalla(e) *Policy*, sia formulato un *Data e Process Management Plan*, che non solo definisce l'organizzazione dettagliata dei processi della gestione dei contenuti digitali lungo il loro ciclo vitale, ma anche la gestione della configurazione degli oggetti digitali rispetto ai cambiamenti dei sistemi in cui essi vengono ospitati e preservati nel tempo.

Definire un *Data e Process Management Plan* potrebbe essere utile per cogliere e formalizzare gli aspetti chiave di un *Trusted Digital Repository*, un Archivio digitale che offre l'accesso affidabile a *lungo termine* alle proprie risorse digitali.

Raggiungere i livelli di un *Trusted Digital Repository* potrebbe dar luogo al coinvolgimento dell'Archivio Istituzionale nel processo di certificazione della sua qualità. La qualità deve riguardare i livelli della *performance* di tutte le fasi del processo di formazione, acquisizione, analisi, elaborazione, valutazione, preservazione, disseminazione e fruizione dei dati (contenuti) relativi alla produzione scientifica dell'Ateneo depositata nell'Archivio il quale deve essere in grado di competere con altri Archivi a livello nazionale e internazionale con costi contenuti.

PARTE II

Indagine su alcuni strumenti a supporto degli Archivi digitali di ricerca

Nella presente Parte della Tesi il lettore è introdotto alla conoscenza di:

1) un'indagine su alcune piattaforme nazionali e internazionali a sostegno della comunicazione del sapere in rete e del Libero Accesso ai contenuti di ricerca;

2) alcuni aspetti a supporto del profilo tecnico di Archivi digitali di ricerca. Il medesimo profilo riguarda l'allestimento di un *Toolkit* (attraverso i riferimenti ad alcuni approcci teorico-pratici, ad una serie di rilevanti progetti e buone pratiche) che comprende i seguenti aspetti: "Panorama di piattaforme" – "Metadati a supporto dei contenuti digitali"- "Controllo di autorità dei valori di metadati" – "Associazione di metadati con valori semantici" – "Conservazione a lungo termine" – "Digitalizzazione" – "Paradigmi di ricerca e navigazione".

3) alcuni approcci per lo sviluppo di uno schema di metadati a supporto descrittivo dei contenuti di un Archivio Istituzionale. Lo schema proposto è sviluppato prendendo in considerazione le "Linee guida per la creazione e la gestione di metadati nei *Repository Istituzionali* (Italia, CRUI, 2012) [V.sito] e le Raccomandazioni "LODE-BD" (*Linked Open Data – enabled Bibliographical Data*) (AIMS, 2012) [V.sito].

CAPITOLO 3

Un'indagine su alcune piattaforme nazionali e internazionali a sostegno della comunicazione del sapere in rete e del Libero Accesso ai contenuti di ricerca

*Digital Repositories for Research output
are on their way to become a permanent part
of the Scholarly Communication
and documentation infrastructure*
[Maurits van der Graaf, 2008, p.88]

3.0. Introduzione

Il presente Capitolo indaga lo stato dell'arte di rappresentazione e di ricerca dei contenuti digitali relativi ad un campione di piattaforme nazionali e internazionali a sostegno della comunicazione del sapere in rete e del Libero Accesso ai contenuti di ricerca. Le medesime piattaforme comprendono:

a) una serie di Archivi (*Repository*) rivelati dal servizio online "*Ranking Web of Repositories*" [V. sito];

piattaforme dei contenuti di ricerca promosse dai progetti "*VOA3R*" (*Virtual Open Access Repository*) e "*The European Library*" ovvero il database TEL - il catalogo unico delle biblioteche nazionali europee.

In riferimento al campione dell'indagine che riguarda 23 piattaforme di gestione dell'informazione che disseminano la conoscenza sui contenuti di ricerca prodotti in diversi paesi, il Capitolo si pone l'obiettivo di esaminare le seguenti questioni:

- 1) quali sono le tipologie di contenuti digitali ospitati nelle medesime piattaforme?
- 2) quali sono le strategie per la gestione, conservazione, ricerca, rappresentazione dei contenuti?
- 3) su quali meccanismi si poggia la comunicazione dei contenuti digitali entro e fuori le piattaforme di riferimento? (es. strumenti di *alerting* che informano l'utente sui contenuti recentemente depositati; *linking* tra i contenuti entro e fuori le piattaforme di riferimento; le statistiche d'uso; i meccanismi di *Social Networks*, *Feedback*; spazi per valutazioni, commenti, *Blog*);

- 4) quali potrebbero essere i prossimi passi per rafforzare il valore degli Archivi Istituzionali in termini di miglioramento dei meccanismi di ricerca, di rappresentazione, di commenti e di interazione dei contenuti di oggetti digitali?

Le 23 piattaforme considerate per l'indagine - tutte compatibili con il protocollo dell'interoperabilità per gli Archivi aperti *Open Archive Initiative Protocol for Metadata Harvesting* - sono distribuite secondo le seguenti categorie:

1. 20 sono *Repository* (tra Istituzionali, Disciplinari, e un fotografico) di 11 diversi paesi creati a sostegno della promozione del materiale di ricerca *Open Access*. Ai medesimi Archivi viene attribuito il *ranking* dal servizio online "*Ranking Web of Repositories*";
2. *OpenAIRE Orphan Repository*: un'infrastruttura digitale per l'aggregazione dei lavori di ricerca prodotti entro il Settimo Programma Quadro. In questo Archivio vengono depositati i materiali di ricerca prodotti dai ricercatori di quelle Istituzioni che non possiedono ancora il proprio *Repository*. Il materiale del *Repository* deve essere accessibile in modalità *Open Access*, immediatamente o dopo un periodo di embargo (6-12 mesi);
3. *The European Library*: un Portale che dissemina la conoscenza sulle risorse bibliografiche di 48 Biblioteche digitali Europee (incluso quelle del materiale di ricerca).

Riferendosi alle piattaforme di cui sopra, il presente Capitolo mira ad indagare le seguenti questioni (Tabella 7) :

1.	Le piattaforme indagate sono allestite con le sezioni a supporto informativo che riguardano le questioni <i>Open Access</i> ?
2.	Gli Archivi indagati rendono esplicita l'informazione sulle proprie Politiche (<i>Policy</i>), sulle eventuali questioni di <i>Copyright</i> , sui metadati e sui formati impiegati a supporto degli oggetti digitali che ospitano?
3.	Quali sono le funzionalità di ricerca e navigazione dei contenuti digitali? Le medesime funzionalità sono affiancate dagli schemi di organizzazione della conoscenza (<i>Knowledge Organization Systems</i>) come <i>Thesauri</i> , <i>Classificazioni</i> e altri vocabolari controllati, delle tecnologie del Web Semantico come, ad esempio, <i>Linked Data</i> ?
4.	Quali sono le modalità di presentazione (<i>layout</i>) riferita ai metadati (indici descrittivi, schede bibliografiche) e un eventuale testo pieno (<i>full-text</i>) degli oggetti digitali rilevati dopo la procedura di ricerca e navigazione?
5.	Quali sono gli eventuali meccanismi a supporto dell'interazione (comunicazione) tra contenuti di <i>Repository</i> (es. accordi di cooperazione tra <i>Repository</i> ; collegamenti/ <i>linking</i> a supporto della condivisione dei contenuti; moduli per commenti, valutazioni e condivisione delle informazioni in riferimento agli oggetti rilevati, alle attività e funzionalità del <i>Repository</i>)?

Tabella 7. Passi d'indagine sulle piattaforme di ricerca dei contenuti digitali

In riferimento ai passi d'indagine appena dichiarati:

- 1) per ogni piattaforma verrà allestita una scheda compilativa;
- 2) nella sezione "Risultati d'indagine" verranno riportati le caratteristiche degli Archivi che risultano efficaci nel recupero e nella rappresentazione dei contenuti, nonché saranno evidenziati i canali comunicativi dei contenuti che potenziano l'interazione tra gli oggetti digitali entro e fuori le piattaforme di ricerca.

L'eventuale rilevamento dei pregi e delle mancanze nell'allestimento delle funzionalità delle piattaforme potrebbe fungere da indicatore di livello di qualità delle medesime piattaforme, che devono essere concepite:

- a) non solo come strumenti per la ricerca dei contenuti, ma anche
- b) come le infrastrutture digitali capaci di assicurare che loro servizi siano trasparenti e affidabili e, quindi, possono essere di beneficio per le istituzioni pubbliche e private, così come per i singoli individui (“*capable of ensuring that their services are transparent and trustworthy, and therefore beneficial to public and private organizations, as well as individuals*” [Duranti, 2012, p. 2].

L’indagine che verrà trattata nel presente Capitolo è stata ispirata dalle informazioni fornite da:

1. il documento “*The European Repository Landscape: Inventory Study into the Present Type and Level of OAI-Compliant Digital Repository Activities in the EU*”, il documento sull’indagine dello stato dell’arte dei *Repository* Istituzionali nei 27 paesi europei che dichiara i seguenti cinque passi d’indagine:
 “DRIVER¹ has identified five specific, complex and long-term issues which are essential to either the establishment, development or sustainability of a Digital Repository:
 1) the business of Digital Repositories,
 2) stimuli for depositing materials into repositories,
 3) intellectual property rights,
 4) data curation, and long- term preservation [...] Special attention is being paid to
 5) the interoperability of repositories to enhance the exchange of data in Repositories”
 [Maurits van der Graaf, Kwame van Eijndhoven, 2008, pp.11, 12];
2. la dodicesima (2013) edizione di “*Ranking Web of Repositories*”² [V. sito; Aguillo et al., 2010] che attribuisce ad ogni *Repository* la sua posizione tra altri Archivi a livello nazionale, europeo ed internazionale;
3. i dati di censimento della *Directory* di autorità di *Open Access Repositories* DOAR [V. sito];
4. il Portale informativo *OpenAIRE (Open Access Infrastructure for Research in Europe)* [V. sito] che promuove *online* diversi riferimenti utili per supportare l’implementazione e l’interoperabilità delle infrastrutture di ricerca *Open Access*;
5. l’articolo “*Linked Data Services for Theses and Dissertations*” [Johnson, Boock, 2012], in cui viene presentato un possibile impiego della tecnologia *Linked Data* per il potenziamento dell’interoperabilità dei contenuti digitali di ricerca [Baker, Tomas (b), 2012].

3.1. Campionamento: selezione delle piattaforme

In accordo con quanto esposto nel paragrafo precedente, nella presente indagine conoscitiva verranno analizzate 23 infrastrutture dei contenuti digitali di ricerca localizzate in diversi paesi. In Tabella 8 (a, b) le medesime infrastrutture saranno descritte ed enumerate con i riferimenti della loro posizione secondo il citato servizio “*Ranking Web of Repositories*”, dei paesi e dei siti web che identificano il loro collocamento.

1.	<p>Il <i>Repository</i> PHAIDRA (l’Università di Vienna, Austria) è sviluppato sulla personalizzazione locale del <i>Software Open Source</i> per gli Archivi aperti - <i>Fedora</i>. Durante il periodo di ricerca all’estero si è avuta la possibilità di seguire le attività a supporto della gestione di questo <i>Repository</i> personalmente.</p> <p>La piattaforma di PHAIDRA:</p> <ul style="list-style-type: none"> • rappresenta una vetrina di tutta la produzione digitale/digitalizzata di ricerca, didattica, e quella amministrativa dell’Università di Vienna;
----	--

¹ Il progetto DRIVER ha preceduto il progetto OpenAIRE nella creazione di base di conoscenza per l’*European Research Area*.

² Questo servizio di *Ranking* è un’iniziativa in corso dal 2008 del *Cybermetrics Lab* dello spagnolo *Consejo Superior de Investigaciones Científicas (CSIC)*. Dai numeri di *Repository* censiti dal servizio appare che vi è un crescente ed attivo interesse nell’implementazione di *Repository* digitali di ricerca in tutto il mondo.

	<ul style="list-style-type: none"> • offre la base tecnologica per l'Archivio Istituzionale del medesimo Ateneo "<i>u:scholar</i>"; • sostiene la gestione e la disseminazione di oggetti depositati appoggiati sui meccanismi per la conservazione a lungo termine; • incorpora le funzionalità personalizzate avanzate per il deposito e la rappresentazione dei contenuti come "<i>Book viewer</i>", "<i>Book exporter</i>", "<i>PHAIDRA exporter</i>"; • offre i diversificati livelli d'accesso per l'utenza, per garantire la tutela del Diritto d'autore; • ha dato l'origine allo sviluppo di una rete di collaborazione "<i>Phaidra International</i>" [V.sito] tra <i>Repository</i> sviluppati in diversi paesi sulla stessa infrastruttura tecnologica.
2.	18 <i>Repository</i> (17 Istituzionali ed 1 Disciplinare " <i>Europe PubMed Central Subject Repository</i> ") implementati in 10 paesi.
3.	L'Archivio aperto <i>Open Access</i> " <i>OpenAIRE Orphan Repository</i> ", che è un'infrastruttura digitale sviluppata entro il progetto <i>OpenAIRE</i> , finanziato dal Settimo Programma Quadro FP7 della CE. Il progetto <i>OpenAIRE</i> ha la <i>mission</i> di facilitare l'accesso a tutta la produzione scientifica dello Spazio Europeo della Ricerca (ERA). Nell'" <i>OpenAIRE Orphan Repository</i> " vengono depositati i lavori di ricerca prodotti entro il FP7 da quei ricercatori cui Istituzioni ancora non possiedono un proprio Archivio Istituzionale o Disciplinare.
4.	L'Archivio digitale fotografico <i>Open Access MDC (Digital Memory of Catalunya)</i> [V.sito] identificato attraverso il citato servizio di censimento degli Archivi aperti <i>Open Access - DOAR</i> . L'Archivio MDC rappresenta un potente servizio per la ricerca e la promozione online del patrimonio fotografico di <i>Catalunya</i> (Spagna).
5.	Il Portale " <i>The European Library</i> " (Olanda) rappresenta le funzionalità avanzate per la navigazione e la ricerca semantica delle risorse bibliografiche provenienti da 48 Biblioteche nazionali europee. La ricerca dei contenuti disseminati tramite il Portale è affiancata da alcuni strumenti dell' <i>Authority control</i> per la disambiguazione dei nomi e dalla tecnologia <i>Linked Data</i> a supporto della connessione dei contenuti della lo stesso campo semantico.

Tabella 8 (a): Introduzione delle infrastrutture di ricerca scelte come oggetto di indagine

In Tabella 8 (b) viene presentata una lista enumerativa di *Repository Open Access* del materiale di ricerca – sviluppati in 11 paesi a livello nazionale e internazionale. Ogni *Repository* sarà affiancato dalla posizione quale gli è stata attribuita dal servizio "*Ranking Web of Repositories*", che censisce più di 1200 Archivi secondo una combinazione dei seguenti criteri:

1. **Dimensione (S)**. Numero di pagine trovate sul motore di ricerca GOOGLE (peso 10%).
2. **Visibilità (V)**. Il risultato della radice quadrata nel numero totale dei *link* esterni (*inlinks*) combinato con il numero dei domini di riferimento, secondo i due maggiori Provider di dati - MAJESTIC SEO and AHREFS (peso 50%).
3. **Rich files**. La somma dei seguenti formati di file: pdf, doc+docx, ppt+pptx and ps+eps (peso 10%)
4. **Scholar**. Numero totale dei documenti trovati tramite GOOGLE SCHOLAR nel periodo 2008-2012 (peso 30%).

I *Repository* - oggetto di indagine - verranno presentati in vista di 6 categorie di Software su cui sono sviluppati (Fedora, PMC, Invenio, DSpace, HAL, EPrints).

	Posizione nella classifica “Ranking Web of Repositories”		Riferimenti web di <i>Repository</i>	Riferimenti web di Software
	<i>World Ranking</i>	<i>Continental/ Country Ranking</i>		
1.	1151	470/ 6  Austria	PHAIDRA < http://bibliothek.univie.ac.at/phaidra.html >	Fedora Open Source Commons: < https://wiki.duraspace.org/display/FEDORA34/Fedora+Digital+Object+Model >
2.	6	1 / 1  Inghilterra	Europe PubMedCentral: < http://europepmc.org >	PMC: < http://www.pmcsoftware.com/ >
3.	8	2 / 1  Svizzera	CERN Document Server: < http://cds.cern.ch/ >	Invenio: < http://invenio-software.org/# >
4.	-	-/-  Svizzera	OpenAIRE Orphan Repository : < http://openaire.cern.ch/ >	Invenio
5.	10	3/ 1  Spagna	Dipòsit Digital de Documents de la UAB: < http://ddd.uab.cat/?ln=en >	Invenio
6.	14	4 / 6  Stati Uniti	MIT DSpace del Massachusetts Institute of Technology: < http://dspace.mit.edu/ >	DSpace : < http://www.dspace.org/ >
7.	173	78/13  Inghilterra	Cambridge University DSpace: < http://www.dspace.cam.ac.uk >	DSpace
8.	173	79 / 12  Inghilterra	University of Edinburgh Research Archive: < http://www.era.lib.ed.ac.uk/ >	DSpace
9.	71	18 / 18  Stati Uniti	Oregon State University ScholarsArchive@OSU: < http://ir.library.oregonstate.edu/xmlui/ >	DSpace

10.	47	18 / 2  Belgio	Université de Liège Open Repository and Bibliography ORBi : < http://orbi.ulg.ac.be >	DSpace
11.	186	87 / 5  Italia	Cadmus European University Institute Institutional Repository: < http://www.eui.eu/Research/EUIPublications/AcademicPublications/AboutCadmus.aspx >	DSpace
12.	17	6 / 2  Spagna	Universitat Politècnica de Catalunya UPCCommons: < http://upcommons.upc.edu/?locale=en >	DSpace
13.	33	11 / 1  Portogallo	Minho's Institutional Repository: < http://repositorium.sdu.m.uminho.pt/handle/1822/6177 >	OAI Extended Addon Dspace 1.8.2: < http://projecto.rcaap.pt/index.php/lang-en/consultar-recursos-de-apoio/remository?func=startdown&id=353 >
14.	9	2 / 1  Francia	HAL Institut National de Recherche en Informatique et en Automatique Archive Ouverte: < http://hal.inria.fr/ >	HAL: < http://www.hal-software.it/it/ >
15.	28	10 / 2  Inghilterra	University of Southampton EPrints: < http://eprints.soton.ac.uk/ >	EPrints: < www.eprints.org >
16.	127	59 / 9  Inghilterra	University College London ePrints: < http://eprints.ucl.ac.uk/ >	EPrints
17.	52	21 / 4  Paesi Bassi	Repository of the University of Twente Publications: < http://doc.utwente.nl/ >	EPrints
18.	22	7 / 1  Germania	Munich Personal RePEc Archive: < http://mpra.ub.uni-muenchen.de/ >	EPrints

Altri due *Repository* sono sviluppati in collegamento con i moduli di *data governance* a supporto della gestione dei contenuti e processi degli Atenei:

1) SURplus (Sistema Università Ricerca plus) - *Open Archive* - basato sul *Software DSpace* - sviluppato dal Consorzio Interuniversitario Lombardo per l'Elaborazione Automatica CILEA;

2) U-GOV – <i>Ricerca</i> , sviluppato dal Consorzio Interuniversitario per il Calcolo Automatico dell'Italia CINECA, trattati nella PARTE I (Capitolo 2).				
19.	160	70/ 4  Italia	Institutional Archive of Research, AIR (sviluppato sul SURplus Open Archive): < http://www.unimi.it/rice/rca/air/20410.htm >	DSpace- modulo SURplus <i>Open Archive</i>
20.	892	0 / 22  Italia	Publication Open Repository Torino PORTO (collegato col modulo U-GOV_ <i>Ricerca</i>): < http://porto.polito.it/ >	EPrints – collegamento col modulo UGOV-<i>Ricerca</i>

Tabella 8 (b). Introduzione delle infrastrutture di ricerca scelte come oggetto di indagine

Secondo la classifica data dal servizio “*Ranking Web of Repositories*”, nell’implementazione dei *Repository* Istituzionali l’Italia sembra confermare una dinamica comune a tutti i paesi maggiormente sviluppati, però, ad un esame più attento delle posizioni attribuite nel quadro generale, emergono differenze sostanziali rispetto ad altri Paesi più avanzati in questo campo, come la Svizzera, la Gran Bretagna, i Paesi Bassi, la Germania e la Francia, solo per restare in un contesto europeo.

Tutti i *Repository* presentati in Tabella 8 (b) fanno parte dell’elenco degli Atenei aderenti alla “Dichiarazione di Berlino sull’Accesso Aperto”, con relativa data di registrazione della loro firma [V.sito: “Conferenza di Berlino”].

I risultati d’indagine di tutte le piattaforme introdotte in Tabella 8 (a,b) saranno allestiti attraverso le schede compilative in cui gli aspetti di osservazione saranno enumerati come verrà presentato nel paragrafo successivo. La scelta del metodo delle schede è stato ritenuto privilegiato anche perché la presentazione visuale e sintetica che esse propongono potrebbe avere il pregio di essere immediatamente comprensibile.

3.2. Domande d’indagine

Per mezzo delle schede compilative - che saranno allestite per ogni infrastruttura d’indagine – saranno messe in luce gli aspetti di cui in Tabella 9.

Prima di iniziare l’allestimento delle schede con le informazioni che descrivono le piattaforme rappresentanti l’oggetto d’indagine, si è cercato di utilizzare i contatti diretti dei *Repository* per spedirvi via posta elettronica alcune domande al fine di ottenere i dati specifici, o più completi, uniformi e confrontabili. Intanto, tale scelta di operare non si è dimostrata fortunata, in quanto non si otteneva alcun *feed-back*. In questa prospettiva si è scelto, per quanto è stato possibile, di procedere alla descrizione degli aspetti – su cui si è basata l’indagine - con le informazioni ricavabili da un diretto esame delle piattaforme online dei *Repository*, integrandone con le risorse disponibili in rete.

L’indisponibilità e/o l’incompletezza dei dati ricavati da alcune piattaforme d’indagine può essere presa in considerazione come un punto da migliorare, in quanto la disponibilità e la completezza delle informazioni relativamente alle funzionalità di una qualsiasi infrastruttura di gestione dei contenuti rileva anche il livello della sua qualità nell’erogazione dei servizi all’utenza.

Gli aspetti valutativi sono stati strutturati in cinque sezioni articolate come segue (Tabella 9):

I.	Presentazione del <i>Repository</i> (dominio web, interfaccia, Software, lingua/e)
	Finalità
	Livelli di accesso
	Policy del <i>Repository</i> (trovato +/- non trovato -)
	Dichiarazioni sull' <i>Open Access</i> (trovato +/- non trovato -)
	Politiche di editori, <i>COPYRIGHT</i> (trovato +/- non trovato -)
	Contenuti di ricerca rilevati
II.	Aspetti tecnici a supporto di presentazione di oggetti depositati
	Dichiarazioni sui Metadati a supporto dei contenuti (trovato +/- non trovato -)
	Dichiarazioni sui Formati (trovato +/- non trovato -)
	Assegnazione di indirizzi/identificatori persistenti ai contenuti depositati
	Memorizzazione di <i>stream</i> di video e audio
	Statistiche d'uso (accesso, visualizzazione, download)
	Funzionalità di <i>upload</i>
	Funzionalità di <i>alerting</i> e <i>feed RSS/Atom</i> per aggiornare sul deposito di nuovi contenuti
III.	Ricerca, navigazione di contenuti
	Funzionalità di ricerca e di navigazione/ <i>Browsing</i>
	Semantica e autorità delle voci di navigazione. Eventuali collegamenti (<i>linking</i>) con <i>Knowledge Organization Systems</i> (es. Classificazioni, Thesaurus) caricati sul Server del <i>Repository</i>
IV.	Presentazione (visualizzazione) dei risultati di ricerca e di navigazione
	Elenco di contenuti recuperati tramite le funzionalità di ricerca e navigazione/ <i>Browsing</i>
	Schede descrittive (record di metadati) per gli oggetti rilevati
	Visualizzazione del contenuto dell'oggetto rilevato
V.	Comunicazione mediata dal <i>Repository</i> in collegamento con altri sistemi, moduli
	Moduli di Feedback, <i>Forum</i> , <i>Blog</i> per la comunicazione attiva con l'utenza
	Moduli per la valutazione (<i>Rating</i>) di contenuti
	Moduli per i commenti (<i>Comments</i> , <i>Discussions</i>) critici collaborativi su contenuti
	Collegamenti con i moduli di <i>Social Networks</i>
	Collegamenti con i contenuti di altri sistemi

Tabella 9. Aspetti valutativi per indagare le piattaforme di *Repository* e Portali

L'indagine svolta non ha la pretesa di avere un carattere esaustivo, bensì di focalizzare le caratteristiche comuni e quelli specifici che possono contribuire ed apportare valore aggiunto a tutte le infrastrutture digitali di ricerca.

3.3. Schede compilative

Nelle seguenti pagine saranno presentate le schede compilative relative all'indagine di 23 piattaforme nazionali e internazionali a sostegno della comunicazione del sapere in rete e del Libero Accesso ai contenuti di ricerca, in corrispondenza alle domande d'indagine definite nel paragrafo precedente.

PHAIDRA (Permanent Hosting, Archiving and Indexing of Digital Resources and Assets). Archivio (*Repository*) digitale dell'Università di Vienna (Austria): <http://bibliothek.univie.ac.at/phaidra.html>

I. Presentazione dell'Archivio/Repository:

The screenshot shows the Phaidra website interface. At the top, there's a header with the university logo and navigation tabs: Recherche, Services, Bibliotheken, News, Mein Konto. Below this, there's a main content area with a large 'PHAIDRA' logo and a description of the system. On the left, there's a sidebar with various services like 'Öffnungszeiten', 'Bibliotheksausweis', 'Entlehnung', etc. At the bottom, there's a search bar and a list of languages: Italiano, Tedesco, Inglese, Serbo.

2. Accessi per la ricerca di contenuti per quattro tipi di utenti: "Administrator", "Super-User", "Owner", "Anonyme User".

È possibile: a) un accesso libero a livello globale in rete per la ricerca e la consultazione degli oggetti (full-text e/o Abstract) depositati e delle schede di metadati; b) scaricare full-text se tale non è impedito dal Copyright, dal periodo di embargo, dalle restrizioni d'accesso ai determinati gruppi di utenti.

Contenuti di ricerca:

1. **Oggetti depositati:** e-books (libri digitali/digitalizzati) e/o loro capitoli, monografie, multimedia (audio, video), collezioni di immagini, articoli, rapporti tecnici, interventi in convegni e conferenze, working papers, tesi di dottorato. Materiale didattico e amministrativo.
 2. **Schede descrittive** (record di metadati) per la rappresentazione di oggetti depositati.

II. Aspetti tecnici per la presentazione di oggetti depositati:

1. Creazione di "eBooks" e collezioni di immagini
2. Memorizzazione di stream di video e audio
3. "Imageviewer": la possibilità di visualizzare l'immagine in diverse prospettive di grandezza
4. "Bulk-Uploads": la possibilità di caricare mega-file con un solo upload
5. "Book Viewer", "Book Importer", "Phaidra importer" (per collezioni)
6. Linking di diversi oggetti tra di loro
7. Indirizzi Persistenti per ogni oggetto depositato
8. "Latest additions": un riepilogo regolarmente aggiornato sui nuovi oggetti depositati tramite le funzioni di Alerting, feed RSS/Atom
9. Presentazione di oggetti in diverse versioni
10. Mailing list
11. È previsto l'impiego sul Server di Knowledge Organization Systems del GETTY Thesaurus, Dewey Decimal Classification disponibili in modalità Linked Data

POLICY del Repository

Dichiarazioni sull'OPEN ACCESS

Politiche di editori, COPYRIGHT

Dichiarazioni su METADATI

Dichiarazioni su FORMATI

Presentazione del contenuto dell'oggetto rilevato e delle sue parti:

The screenshot shows a document viewer interface. It displays a scanned page of text, likely a historical document, with a search bar and navigation controls. The text on the page is in German and appears to be a letter or a document from the 18th century.

V. Comunicazione mediata dal Repository in collegamento con altri sistemi, moduli:

1. L'interazione diretta dei contenuti di PHAIDRA con i contenuti di:
 - a) Repository del materiale di ricerca dell'Ateneo U:SCOLAR
 - b) servizi dell'Ateneo: Anagrafe di Ricerca RAD, "E-Learning Plattform";
 - c) Confederazione di Open Access Repositories COAR; Archivi aggregatori "Europeana" (collezioni culturali europee), "OpenAIRE" (collezioni del materiale di ricerca europee),



2. Cooperazione con: il progetto Europeo Tempus; 14 istituzioni di 5 diverse paesi che hanno adottato la piattaforma "Phaidra", formando assieme una rete Phaidra International: "Phaidra.org"¹



III. Navigazione e ricerca dei contenuti:

The screenshot shows a search interface with various filters and options. It includes sections for 'Per struttura', 'Per classificazione', 'Per tipo di oggetto', and 'Per lingua'. Each section has a dropdown menu to select different search criteria.

L'impiego di una serie di Classificazioni per la navigazione a faccette (o multidimensionale):

The screenshot shows a dropdown menu for 'Classificazioni utilizzate'. It lists several classification schemes, including 'Physics and Astronomy Classification Scheme', '30.00.00: ATOMIC AND MOLECULAR PHYSICS', '34.00.00: Atomic and molecular collision processes and interactions (for atomic, mol...', '34.50.-s: Scattering of atoms and molecules', and '34.50.0b: Electronic excitation and ionization of molecules'.

IV. Presentazione dei risultati di ricerca e di navigazione:

The screenshot shows a search results page with a list of items. Each item has a small thumbnail image and a title. The items are related to historical documents and books.

Presentazione della scheda descrittiva (registrazione attraverso i metadati) dell'oggetto rilevato:

The screenshot shows a detailed metadata record for a specific document. It includes fields for 'Descrizione (de)', 'Autori', 'Classificazioni', and 'Fa parte della collezione'. The record provides comprehensive information about the document, including its title, author, and classification.

¹ <http://phaidra.org/>



Europe PubMed Central Subject Repository. Consorzio *Europe PMC Funders' Groupe* la rete degli Archivi “*PMC International*”² (Inghilterra):

[<http://europepmc.org/>](http://europepmc.org/)

I. Presentazione dell'Archivio/Repository:

Software: PMC

Lingua di presentazione: Inglese

Finalità: lanciato come “*UK PubMed Central*” (UKPMC) nel 2007, cambiato per “*Europe PubMed Central*” nel 2012, il *Repository* ha l’obiettivo di raccogliere, preservare e disseminare i contenuti di ricerca del campo biomedico attraverso la rete degli Archivi *PMCI* e globalmente in retetramite l’Internet

Livelli d’accesso:

1. Un **accesso autorizzato** tramite “*Login*” per il deposito di contenuti e per l’uso delle funzionalità avanzate del sistema come *linking* contestuale, visualizzazione delle citazioni, statistiche d’uso.
2. Un **accesso libero** per la ricerca e la consultazione di oggetti (*full-text* e/o *Abstract*) depositati e schede di metadati. È possibile scaricare *full-text* se tale non è impedito dal *Copyright* e/o dal periodo di embargo.

Contenuti di ricerca:

1. **Oggetti depositati** articoli, patenti, linee guida, dati di ricerca. Tra 2.2 milioni di contenuti depositati in *full-text*, più di 400.00 sono in modalità d’accesso *Open Access*
2. **Schede descrittive** (record di metadati) per la rappresentazione di oggetti depositati.

II. Aspetti tecnici per la presentazione di oggetti depositati:

1. **Indirizzi Persistenti** per ogni oggetto depositato.
2. **Statistiche d’uso** (accesso, visualizzazione, download)
3. Visualizzazione delle **citazioni** per l’articolo rilevato e esportazione delle citazioni tramite *Citation Index Services*
4. Un riepilogo regolarmente aggiornato sui nuovi “*Recent*” e più consultati “*Popular*” oggetti depositati tramite le funzioni *Alerting, feed RSS/Atom*
5. Consultazione “*Copyright and License information*” per ogni oggetto depositato

III. Navigazione e ricerca dei contenuti:

1. **Navigazione contestuale fra gli oggetti semanticamente correlati** attraverso le voci di ricerca: “Autore”, “Titolo”, “Affiliazione”, “Soggetto” (normalizzato con le voci dei vocabolari controllati semantici), “Riviste”, “*Granti*”
2. “*Related Articles*”: visualizzazione degli articoli semanticamente correlati all’articolo rilevato.
3. Consultazione delle **banche dati di ricerca** a partire dalle **parole estratte** ed evidenziate nel testo e/o nell’*Abstract* con la tecnica di *Text-Mining*.
4. **Ricerca semantica** dei contenuti tramite la funzione “*Europe PMC labs*”

IV. Presentazione dei risultati di ricerca e di navigazione:

Presentazione della descrizione e del contenuto dell’oggetto rilevato:

V. Comunicazione mediata dal Repository in collegamento con altri sistemi, moduli:

1. Interazione dei contenuti del *Repository* con:
 - a) banche dati (esterne, interne) di dati di ricerca;
 - b) contenuti dai siti di editori;
 - c) valori di vocabolari controllati semantici *MeSH*, *Swissprot*, *Gene Ontology (GO)*, *NCBI's taxonodline Plus*;
2. Funzionalità di *Feedback* e *Blog* per la comunicazione attiva al servizio dell’utente.
3. Moduli per la valutazione (*Rating*) dei contenuti da parte di utenti registrati

POLICY del *Repository*

+ **Dichiarazioni sull’OPEN ACCESS**

+ **Politiche di editori, COPYRIGHT**

+ **Dichiarazioni su METADATI³**

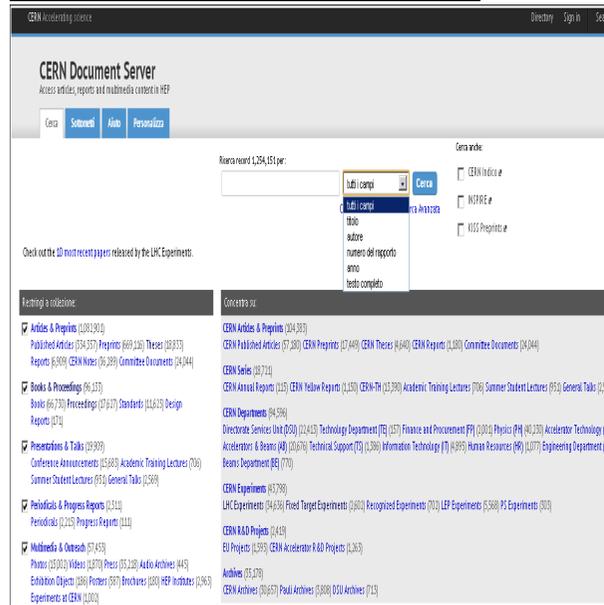
+ **Dichiarazioni su FORMATI** + **trovato** + **non trovato**

² <<http://europepmc.org/Funders/>>; <<http://www.ncbi.nlm.nih.gov/pmc/about/pmc/>>

³ EBI Europe PMC Web Service 3.0.1, Reference Guide, 18 th December 2012, Doc Version 1.4, <http://europepmc.org/docs/EBI_Europe_PMC_Web_Service_301_Reference_14.pdf>



I. Presentazione dell'Archivio/Repository:



Software: *Invenio Open Source*

Lingue di presentazione: Inglese, Italiano, [Български](#) [Català](#) [Deutsch](#) [Ελληνικά](#) [English](#) [Español](#) [Français](#) [Hrvatski](#) [日本語](#) [ქართული](#) [Norsk/Bokmål](#) [Polski](#) [Português](#) [Русский](#) [Slovensky](#) [Svenska](#) [中文\(简\)](#) [中文\(繁\)](#)

Finalità: raccogliere, preservare e disseminare i contenuti di ricerca e il materiale didattico prodotti entro il CERN

Livelli d'accesso: **1. Accessi autorizzati** tramite "Login" per il deposito di contenuti da parte dello staff accademico e dei ricercatori del CERN

2. Un accesso libero per la ricerca e la consultazione di oggetti (*full-text* e/o *Abstract*) depositati e schede di metadati. È possibile scaricare *full-text* se tale non è impedito dal *Copyright* e/o dal periodo di embargo.

POLICY del Repository + **Dichiarazioni sull'OPEN ACCESS** - **Politiche di editori, COPYRIGHT** + **Dichiarazioni su METADATI** + **Dichiarazioni su FORMATI** + **trovato** + **non trovato** -

Contenuti di ricerca:

- Oggetti depositati** (1.300.000): articoli; interventi in convegni e conferenze; tesi; libri (e-book); materiale didattico; multimedia (audio, video, fotografie); collezioni di Archivi personali.
- Schede descrittive** (record di metadati) per la rappresentazione di oggetti depositati.

II. Aspetti tecnici per la presentazione di oggetti depositati:

- "Tuo account":** spazio per la memorizzazione di ricerche/avvisi/download.
- Indirizzi Persistenti** per ogni oggetto depositato.
- Statistiche d'uso** (accesso, visualizzazione, download)
- System-generated citation metrics**
- Un riepilogo regolarmente aggiornato sui nuovi oggetti depositati tramite le funzioni di **Alerting, feed RSS/Atom.**

III. Navigazione e ricerca dei contenuti:



Raffinamento e correlazione dei risultati di ricerca:



IV. Presentazione dei risultati di ricerca e di navigazione:



Presentazione del contenuto dell'oggetto rilevato:



V. Comunicazione mediata dal Repository in collegamento con altri sistemi, moduli:

1. Interazione tra i contenuti delle collezioni CERN con i contenuti delle collezioni interne ed esterne "CERN Indico", "INSPIRE", "KISS Preprints"

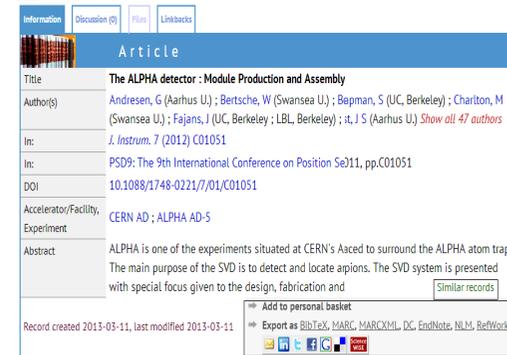
2. Condivisione/esportazione di contenuti in vari formati e in vari *Social networks* a partire dalla scheda descrittiva dell'oggetto depositato

3. Interazione tra contenuti sulla base dei valori del **vocabolario controllato (Thesauro) definito localmente**

4. Funzionalità di **Feedback** per la comunicazione attiva al servizio dell'utente

5. Nella scheda descrittiva dell'oggetto depositato: **a)** modulo "Discussion" per i commenti critici; **b)** contatto **E-Mail** dell'autore depositante

Presentazione della scheda descrittiva (registrazione attraverso i metadati) dell'oggetto rilevato:





I. Presentazione dell'Archivio/Repository:

Software: *Invenio Open Source*

Lingue di presentazione: Inglese

Finalità: raccogliere, preservare e disseminare materiali di ricerca prodotti entro il Settimo Programma Quadro FP7 (2007-2013) da quelle istituzioni che ancora non possiedono un proprio Archivio Istituzionale o disciplinare, o se gli Archivi esistenti non sono compatibili con i requisiti OpenAIRE (*Open Research Infrastructure for Research in Europe*), il progetto che ha lo scopo di sostenere la diffusione dell'accesso aperto in Europa nell'adeguamento alle richieste della Commissione Europea e alle linee guida dell'*European Research Council* sullo stesso tema.

Livelli d'accesso:

1. **Accessi autorizzati** con "Login" tramite il Portale OpenAIRE per il deposito dei contenuti di ricerca prodotti entro il Settimo Programma Quadro FP7
2. **Un accesso libero** per la ricerca e la consultazione dei contenuti (*full-text* e/o *Abstract*) depositati e delle schede di metadati. Lo scarico (*download*) di *full-text* se tale non è impedito dal *Copyright* e/o dal periodo di embargo (6-12 mesi)

Contenuti di ricerca:

1. **Oggetti depositati** (350): articoli; interventi in convegni e conferenze; *e-book*
2. **Schede descrittive** (record di metadati) per la rappresentazione di oggetti depositati.

II. Aspetti tecnici per la presentazione di oggetti depositati:

1. "Tuo account": spazio per la memorizzazione di ricerche/avvisi/download
2. **Statistiche d'uso** (accesso, visualizzazione, download)
3. **System-generated citation metrics**
4. Un riepilogo regolarmente aggiornato sui nuovi oggetti depositati tramite le funzioni di *Alerting, feed RSS/Atom*
5. **File allegati**

III. Navigazione e ricerca dei contenuti:

IV. Presentazione dei risultati di ricerca e di navigazione:

Presentazione della scheda descrittiva (registrazione attraverso i metadati) dell'oggetto rilevato:

Presentazione del contenuto dell'oggetto rilevato:

V. Comunicazione mediata dal Repository in collegamento con altri sistemi, moduli:

1. A partire dalla scheda descrittiva dell'oggetto è possibile:
 - a) condividere/esportare i contenuti in vari formati e in *Social networks*,
 - b) effettuare "Discussion" e "Rating"
2. Interazione di oggetti con:
 - a) set di dati di ricerca,
 - b) informazioni sull'Autore e sul Finanziamento di ricerca.
 Tramite tali associazioni definiti coi link appropriati, una pubblicazione diventa "Enhanced Publication"
3. Funzionalità di *Feedback* per la comunicazione attiva al servizio dell'utente.

POLICY del Repository + Dichiarazioni sull'OPEN ACCESS + Politiche di editori, COPYRIGHT + Dichiarazioni su METADATI¹ + Dichiarazioni su FORMATI + trovato + non trovato -

¹ <<http://www.openaire.eu/en/component/content/article/207>>



Dipòsit Digital de Documents de la UAB (Spagna) : <<http://ddd.uab.cat/?ln=en>>

I. Presentazione dell'Archivio/Repository:



Software: *Invenio Open Source*

Lingue di presentazione: Inglese, Spagnolo, Catalano

Finalità: funzione di un meta-Archivio che raccoglie, gestisce, preserva e diffonde in rete - a partire da un unico punto d'accesso- la conoscenza sul materiale (digitale, digitalizzato) di ricerca, didattico, Archivi personali provenienti da diverse biblioteche digitali dell'Università Autonoma di Barcellona²

Livelli d'accesso: 1. Accessi autorizzati al personale amministrativo dell'Archivio per il deposito dei contenuti dietro un Permesso formale sottoscritto dall'autore con l'indicazione della Licenza d'uso per l'oggetto che verrà depositato

2. Accesso libero a livello globale in rete per la ricerca e la consultazione di oggetti (*full-text* e/o *Abstract*) depositati e delle schede di metadati. È possibile scaricare *full-text* se tale non è impedito dal *Copyright*, dal periodo di embargo, dalle restrizioni d'accesso ai determinati gruppi di utenti.

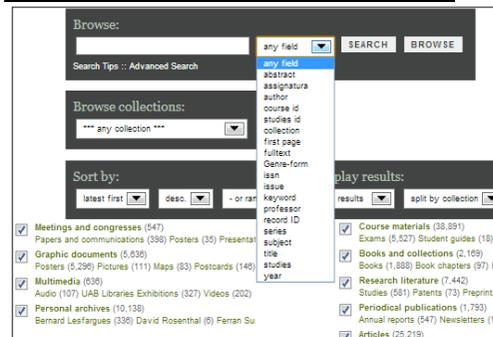
Contenuti di ricerca:

1. **Oggetti depositati** (100.000): articoli; interventi in convegni e conferenze; *e-book*; materiale didattico; multimedia (audio, video, fotografie); materiale grafico; collezioni di Archivi personali
2. **Schede descrittive** (record di metadati) per la rappresentazione di oggetti depositati.

II. Aspetti tecnici per la presentazione di oggetti depositati:

1. **Statistiche d'uso** (accesso, visualizzazione, download) per ogni oggetto
2. Riepilogo regolarmente aggiornato sui **nuovi oggetti depositati** tramite le funzioni di *Alerting*, *feed RSS/Atom*
3. **Consultazione di Licenza d'uso** per ogni oggetto

III. Navigazione e ricerca dei contenuti:



IV. Presentazione dei risultati di ricerca e di navigazione:



Presentazione della scheda descrittiva (registrazione attraverso i metadati) dell'oggetto rilevato:



Presentazione del contenuto dell'oggetto rilevato:



V. Comunicazione mediata dal Repository in collegamento con altri sistemi, moduli:

1. Interazione tra i contenuti delle collezioni di **cinque Archivi Digitali spagnoli**, presentati nella Directory della Biblioteca dell'Ateneo.
2. Connessione con diverse **database esterne di dati/contenuti di ricerca**.
3. Moduli per i commenti critici collaborativi (*Discussions*) e per la valutazione (*Rating*) dei contenuti a partire dalla scheda descrittiva dell'oggetto depositato.
4. Possibilità di **condivisione/export** delle informazioni sugli oggetti depositati tramite i collegamento con i vari *Social networks*

POLICY del Repository + OPEN ACCESS + Politiche di editori, COPYRIGHT + Dichiarazioni su METADATI + Dichiarazioni su FORMATI + trovato + non trovato -

² <<http://www.uab.cat/servlet/Satellite/recursos-d-informacio/diposits-digital-1253513285190.html>>



I. Presentazione dell'Archivio/Repository:

The screenshot shows the DSpace@MIT homepage. It features a search bar at the top left, a 'Browse' section with various categories like 'Communities & Collections', 'By Issue Date', 'Authors', 'Titles', and 'Subjects'. There is also a 'My Account' section with 'Login' and 'Logout' buttons. A 'Links' section lists various research centers and departments. The main content area includes a 'Search' box, a 'Communities' list, and a 'Latest News' section with several articles and statistics.

Software: *DSpace open-source*
Lingue di presentazione: Inglese

Finalità: raccogliere e preservare il materiale di ricerca in formato digitale prodotto entro MIT. Disseminare il medesimo materiale tramite l'Internet

Livelli d'accesso:

1. **Accessi autorizzati** tramite "Login" per il deposito di contenuti da parte dello staff accademico e dei ricercatori del MIT
2. **Un accesso libero** per la ricerca e la consultazione di oggetti (*full-text* e/o *Abstract*) depositati e schede di metadati. È possibile scaricare *full-text* se tale non è impedito dal *Copyright* e/o dal periodo di embargo. Molti oggetti possono essere scaricati solo entro la rete MIT.

Contenuti di ricerca:

1. **Oggetti depositati:** (60.000 tra cui 7.500 in *Open Access*): articoli; interventi in convegni e conferenze; *e-book*; rapporti tecnici, tesi, set di dati di ricerca, multimedia, materiale didattico
2. **Schede descrittive** (record di metadati) per la rappresentazione di oggetti depositati.

II. Aspetti tecnici per la presentazione di oggetti depositati:

1. Riepilogo regolarmente aggiornato sui **nuovi oggetti depositati** tramite le funzioni di *Alerting*, *feed RSS/Atom*
2. **Indirizzi Persistenti** per ogni oggetto depositato.
3. **Statistiche generali d'uso** (accesso, visualizzazione, download)
4. **Licenze d'uso** per ogni oggetto depositato
5. **Visualizzazione di statistiche di download** del materiale di ricerca *Open Access* per continenti:



(Updated: September, 2012)
Global impact: Since its launch in October 2009, articles in the MIT Open Access collection have been downloaded from nearly all countries worldwide.

III. Navigazione e ricerca dei contenuti:

Le voci del "Soggetto" sono normalizzate con la *Classificazione Decimale Dewey (DDC)* disponibile come *Linked Data*

The screenshot shows the search results for 'Communities'. It lists several communities including 'Abdul Latif Jameel Poverty Action Lab (J-PAL)', 'Air Transportation Research', 'Auto-ID Laboratory', and 'Biological Engineering Division'. Each entry includes a brief description and a link to browse the collection.

IV. Presentazione dei risultati di ricerca e di navigazione:

The screenshot shows search results for the article 'Against Talking Linguistic Diversity at "Face Value"'. The author is David Pesetsky from Cambridge University Press. The results also show other related articles and a link to the full text.

Presentazione della scheda descrittiva (registrazione attraverso i metadati) dell'oggetto rilevato:

The screenshot shows the detailed metadata page for the article 'Toward understanding natural language directions'. It includes fields for Title, Author, Department, Issue Date, Abstract, URI, ISSN, Citation, Version, Terms of Use, Detailed Terms, Published At, and Journal. It also provides a link to the full text and a 'Downloadable Full Text' button.

Presentazione del contenuto dell'oggetto rilevato:

The screenshot shows the MIT Open Access Articles page for the article 'Toward understanding natural language directions'. It includes a citation, a table with metadata (As Published, Publisher, Version, Accessed, Citable Link, Terms of Use, Detailed Terms), and a list of authors with their affiliations.

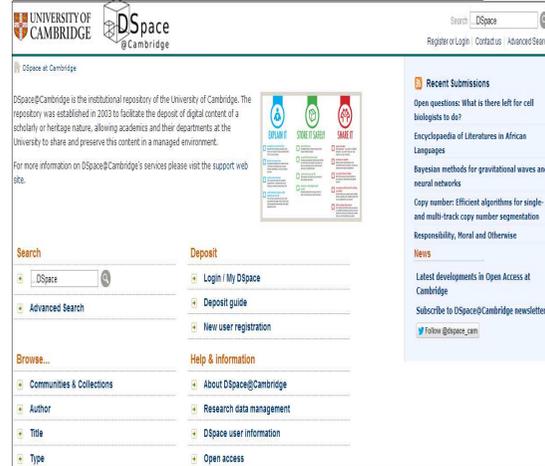
V. Comunicazione mediata dal Repository in collegamento con altri sistemi, moduli:

1. Interazione tra contenuti delle collezioni in base della *Classificazione Decimale Dewey (DDC)*
2. Modulo *Feedback* per le comunicazioni attive con l'utenza.
3. Attraverso l'impiego della DDC pubblicata come *Linked Data*, i contenuti del MIT DSpace possono interagire con altri set di dati affiancati dalla DDC nel



POLICY del Repository + **Dichiarazioni sull'OPEN ACCESS** + **Politiche di editori, COPYRIGHT** + **Dichiarazioni su METADATI** + **Dichiarazioni su FORMATI** + **trovato + non trovato -**

I. Presentazione dell'Archivio/Repository:



Software: *DSpace Open Source*

Lingue di presentazione: Inglese

Finalità: raccogliere e preservare il materiale di ricerca in formato digitale dell'Università di Cambridge. Disseminare la conoscenza sul medesimo materiale tramite l'Internet

Livelli d'accesso:

1. **Un accesso autorizzato** per lo staff accademico e ricercatori dell'Ateneo per il deposito dei propri lavori scientifici

2. **Un accesso libero** per la ricerca e la consultazione di oggetti (*full-text* e/o *Abstract*) depositati e schede di metadati. È possibile scaricare *full-text* se tale non è impedito dal *Copyright* e/o dal periodo di embargo.

POLICY del Repository + **Dichiarazioni sull'OPEN ACCESS**

Contenuti di ricerca:

1. **Oggetti depositati** (200.000): articoli; interventi in convegni e conferenze, relazioni tecniche, tesi, *e-book*, multimedia, materiale didattico, dati di ricerca, materiale amministrativo³
2. **Schede descrittive** (record di metadati) per la rappresentazione di oggetti depositati.

II. Aspetti tecnici per la presentazione di oggetti depositati:

1. Riepilogo regolarmente aggiornato su "*Recent Submissions*" tramite le funzioni di *Alerting, feed RSS/Atom*
2. **Indirizzi Persistenti** per ogni oggetto depositato.
3. Creazione, gestione a accesso ai **dati di ricerca**.
4. **File allegati** agli oggetti depositati
5. **Statistiche d'uso** (accesso, visualizzazione, *download*) per ogni oggetto.
6. **Creazione automatica della citazione** all'oggetto.

III. Navigazione e ricerca dei contenuti:



+ **Politiche di editori, COPYRIGHT** + **Dichiarazioni su METADATI** + **Dichiarazioni su FORMATI** + trovato + non trovato -

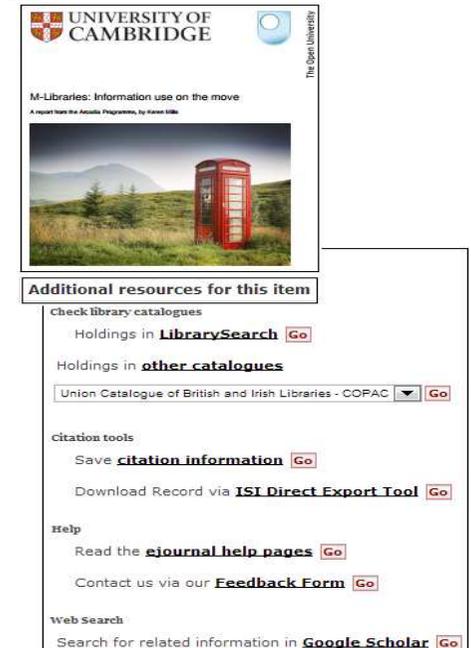
IV. Presentazione dei risultati di ricerca e di navigazione:



Presentazione della scheda descrittiva (registrazione attraverso i metadati) dell'oggetto rilevato:



Presentazione del contenuto dell'oggetto rilevato:



V. Comunicazione mediata dal Repository in collegamento con altri sistemi, moduli:

1. Interazione tra contenuti del *Repository* con contenuti di:

a) cataloghi della biblioteca dell'Università di Cambridge;

b) *Google Scholar*.

2. Funzionalità di *Feedback* per la comunicazione attiva al servizio dell'utente

³ <<http://www.dspace.cam.ac.uk/browse?type=type>>



I. Presentazione dell'Archivio/Repository:

Software: *Dspace Open Source*
Lingue di presentazione: Inglese

Finalità: raccogliere e preservare il materiale di ricerca in formato digitale dell'Università di Edimburgo. Disseminare la conoscenza sul medesimo materiale tramite l'Internet

Livelli d'accesso:

1. Un accesso autorizzato per lo staff accademico e ricercatori dell'Ateneo per il deposito dei propri lavori scientifici
2. Un accesso libero per la ricerca e la consultazione di oggetti (*full-text* e/o *Abstract*) depositati e schede di metadati. È possibile scaricare *full-text* se tale non è impedito dal *Copyright* e/o dal periodo di embargo.

Contenuti di ricerca:

1. Oggetti depositati (6.000): articoli; interventi in convegni e conferenze, tesi, dissertazioni, relazioni tecniche, libri (e-book), dati di ricerca, *working papers*.
2. Schede descrittive (record di metadati) per la rappresentazione di oggetti depositati.

II. Aspetti tecnici per la presentazione di oggetti depositati:

1. Riepilogo regolarmente aggiornato sui "Resent submissions" tramite le funzioni di *email updates (alerting)*, *feed RSS/Atom*
2. Indirizzi Persistenti per ogni oggetto depositato.
3. Consultazione di Licenza d'uso per ogni oggetto depositato
4. Statistiche d'uso (accesso, visualizzazione, download) per ogni collezione e oggetto depositato.

III. Navigazione e ricerca dei contenuti:

IV. Presentazione dei risultati di ricerca e di navigazione:

Issue Date	Title	Author(s)	Supervisor(s)
13-Sep-2011	Form is Function: Faithful Transmission of Information as a Pressure for	Vollmer, Lauren N	Smith, Kenny; Blythe, Richard
Jul-1995	Connections, Learning and Linguistics Structure	Christiansen, Morten H	Chater, Nick; Paul, Schweizer
4-Nov-2007A	Genre Analysis of Michael Moore's Fahrenheit 9/11	Castaldi, Jacopo	Trippes-Lomax, Hugh
2009	The effect of linguistic devices in information presentation messages on	Pietze, Martin; Winterboer, Andi; Moore, Johanna	
2005A	Sociolinguistic study of Chinese lexical change	Su, Yunxuan	Rajah-Carrim, Aaliya
26-Nov-2009	I'm not George! I'm not actually anything! Convergent and divergent identity in a South Durham New-Town.	West, Helen Faye	Maguire, Warren

Presentazione della scheda descrittiva (registrazione attraverso i metadati) dell'oggetto rilevato:

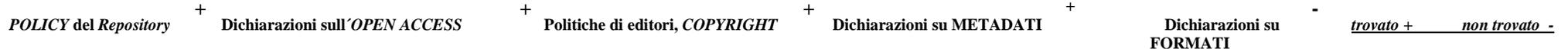
Presentazione del contenuto dell'oggetto rilevato:

SEMANTIC RESPONSIBILITY
Philosophical Explorations, 1, January 2002: 39-58.
 Josefa Toribio
 School of Cognitive and Computing Sciences
 University of Sussex
 Falmer, Brighton BN1 9QH
 England
 UK

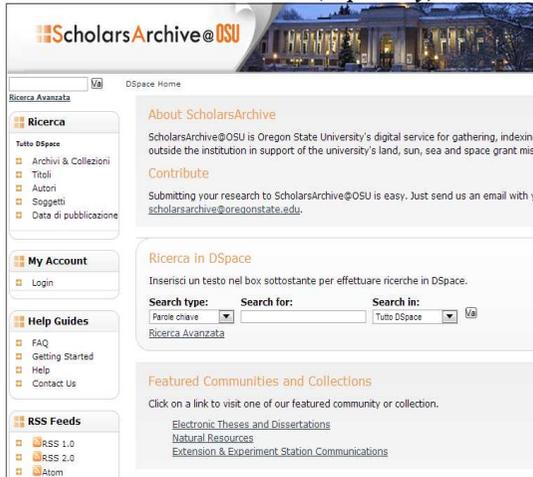
V. Comunicazione mediata dal Repository in collegamento con altri sistemi, moduli:

3. L'interazione tra i contenuti di tutte le collezioni del Repository sulla base delle voci di ricerca e navigazione.

4. Funzionalità di *Feedback* per la comunicazione attiva al servizio dell'utente.



I. Presentazione dell'Archivio (Repository)



Software: *DSpace Open Source*
Lingue di presentazione: Inglese

Finalità: raccogliere e preservare il materiale di ricerca in formato digitale dell'Università di Oregon. Disseminare la conoscenza sul medesimo materiale tramite l'Internet

Livelli d'accesso:

1. **Un accesso autorizzato** per lo staff accademico e ricercatori dell'Ateneo per il deposito dei propri lavori scientifici
2. **Un accesso libero** per la ricerca e la consultazione di oggetti (*full-text* e/o *Abstract*) depositati e schede di metadati. È possibile scaricare *full-text* se tale non è impedito dal *Copyright* e/o dal periodo di embargo.

Contenuti di ricerca:

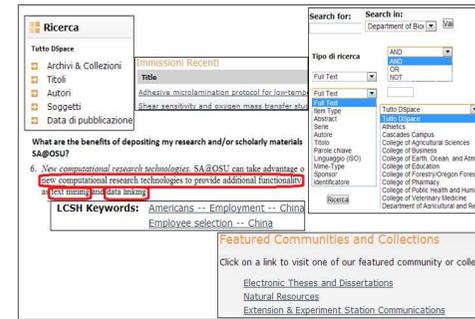
1. **Oggetti depositati:** (35.000): articoli, presentazioni, interventi in convegni e conferenze, relazioni tecniche di dipartimenti, tesi, documenti storici dell'OSU.
2. **Schede descrittive** (record di metadati) per la rappresentazione di oggetti depositati.

II. Aspetti tecnici per la presentazione di oggetti depositati:

1. **Traduzione automatica dei contenuti** esposti nell'interfaccia dell'Archivio in diverse lingue, a seconda del motore di ricerca nazionale da cui si accede
2. **Text-mining** e **data linking** dei contenuti
3. Servizio di ricerca e navigazione "**Linked Data Services for Theses and Dissertations**"¹
4. Riepilogo regolarmente aggiornato sui **nuovi oggetti depositati** tramite le funzioni di **Alerting**, **feed RSS/Atom**.
5. **Indirizzi Persistenti** per ogni oggetto depositato
6. **Statistiche d'uso** (accesso, visualizzazione, download) per ogni oggetto.

III. Navigazione e ricerca dei contenuti:

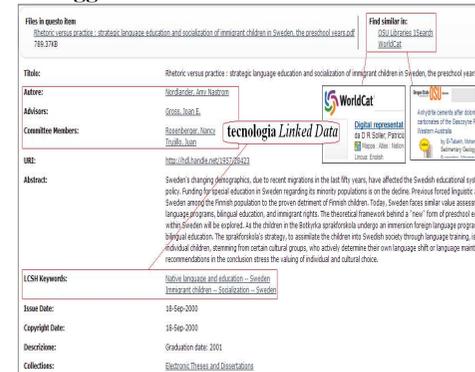
- a) le voci del "Soggetto" sono normalizzate con la *Library of Subject Classification Headings /LCSH* disponibile come *Linked Data*;
- b) le voci del "Autore" sono supportate da link sviluppate con la tecnologia *Linked Data*;
- c) navigazione di contenuti semanticamente correlati nelle collezioni "*OSU Libraries*", nel catalogo digitale delle biblioteche a livello globale "*WorldCat*" (a partire dalla scheda descrittiva)



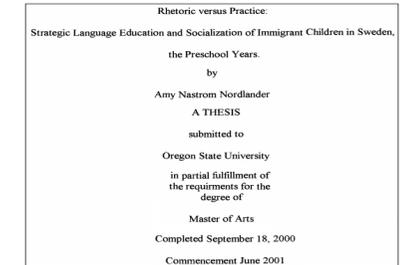
IV. Presentazione dei risultati di ricerca e di navigazione:

Title	Author(s)	Date
Harvest maturity of Cascade and Willamette hops	Sharp, Daniel C. (Daniel Collier)	2013-01-18
A DC-DC converter architecture for low-power, high-resistance	Miller, Brian A.	2013-02-27

Presentazione della scheda descrittiva (registrazione attraverso i metadati) dell'oggetto rilevato:



Presentazione del contenuto dell'oggetto rilevato:

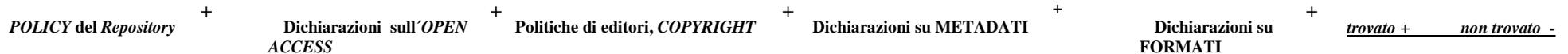


V. Comunicazione mediata dal Repository in collegamento con altri sistemi, moduli:

1. Interazione tra contenuti delle collezioni in base della tecnologia *Linked Data*
2. Interazione tra contenuti del *Repository* con contenuti digitali:
 - a) del catalogo della biblioteca dell'OSU "*OSU Libraries*",
 - b) del Mega-Catalogo delle collezioni delle biblioteche digitali a livello globale *WorldCat*;
 - c) affiancati dalla classificazione LCSH nel



3. Modulo *Feedback* per le comunicazioni attive con l'utenza



¹ Johnson, Thomas, Boock, Michael. *Linked Data Services for Theses and Dissertations*, ETD2012 Conference, Lima, Peru, September 12-14, 2012, <<http://hdl.handle.net/1957/32977>>



Open Repository and Bibliography ORBi¹. Réseau des Bibliothèques. Université de Liège (Belgio): <<http://orbi.ulg.ac.be>>

I. Presentazione dell'Archivio/Repository:

The screenshot shows the ORBi website interface. At the top, there's a navigation bar with 'Home', 'Other OA projects at the ULg', and 'You are here: ORBi > Home'. A search bar is prominently displayed. Below it, there are sections for 'Advanced search' and 'Browse ORBi by' with filters for Author, Title, Issue year, Journal title, Document type, and Discipline. A 'Recently added' section lists items like 'Berlin 10 Open Access Conference' and 'Open Access: the Belgian Ministers commit themselves! The Brussels Declaration on Open Access'. There's also a 'Disciplines in ORBi' section with a list of categories and their counts.

Software: DSpace OpenSource.

Metadati: <<http://orbi.ulg.ac.be/project?id=103>>

Lingue di presentazione: Inglese, Francese

Finalità: raccogliere e preservare il materiale di ricerca digitale/digitalizzato dell'Université de Liège. Disseminare la conoscenza sul medesimo materiale tramite l'Internet.

Livelli d'accesso:

1. Un accesso autorizzato allo staff accademico e ai ricercatori dell'Università di Liège per il deposito del materiale.
2. Un accesso libero per la ricerca e la consultazione di oggetti (full-text e/o Abstract) depositati e schede di metadati. È possibile scaricare full-text se tale non è impedito dal Copyright e/o dal periodo di embargo.

POLICY del Repository + Dichiarazioni sull'OPEN ACCESS + Politiche di editori, COPYRIGHT + Dichiarazioni sui METADATI + Dichiarazioni su FORMATI + trovato + non trovato -

Contenuti di ricerca:

1. Oggetti depositati (95.000 con 56.513 in Open Access, full text): articoli, e-books, reports, working papers, conference proceedings, tesi di dottorato, materiale didattico, presentazioni, rapporti tecnici di dipartimenti, dati di ricerca
2. Schede descrittive (record di metadati) per la rappresentazione di oggetti depositati.

II. Aspetti tecnici per la presentazione di oggetti depositati:

1. "Recently added": riepilogo regolarmente aggiornato sui nuovi oggetti depositati tramite le funzioni di Alerting, feed RSS/Atom.
2. Riferimenti bibliografici di lavori pubblicati dello staff accademico e dei ricercatori dell'Università importati nell'area di lavoro personale MyORBi dai servizi Web of Science e Medline
3. Indirizzi Persistenti per ogni oggetto depositato
4. Statistiche d'uso (accesso, visualizzazione, download) generali

III. Navigazione e ricerca dei contenuti:

The screenshot shows the 'Advanced search' interface. It includes a 'Search index' dropdown set to 'All indexes', a 'Search for:' field, and a 'Document type and subtype' dropdown. There are sections for 'Filters' (Document type, Discipline) and 'Disciplines' with a list of categories like 'Arts & humanities [10049]', 'Archaeology [661]', etc. There are also 'Search' and 'Clear' buttons.

L'impiego della ORBi's Classification basata sulla Intellectual Property & Science ISI Classification per la navigazione a faccette (o multidimensionale).

IV. Presentazione dei risultati di ricerca e di navigazione:

The screenshot shows search results for the query 'linguistics'. The first result is 'Linguistics : An Introduction : traduction d'e RENAVILLE, FRANÇOIS Ulg'. The second result is '"Linguistics : An Introduction" : extraits trad RADFORD, ANDREW; ATKINSON, MARTIN; BRITAIN, DAVID et al Learning material (2002)'. The third result is 'Gestion et référencement des e-journaux à l'ULg : historique, état des lieux & perspectives RENAVILLE, FRANÇOIS Ulg'. The fourth result is 'Comparaison des outils bibliographiques et de Scopus : rapport du groupe de réflexion mixte Interuniversitaire de la Communauté française LERINCKX, DOMINIQUE; DAGUET, MURIEL; RENAVILLE, FRANÇOIS Ulg et'.

Presentazione della scheda descrittiva (registrazione attraverso i metadati) dell'oggetto rilevato:

The screenshot shows the detail page for the document 'Gestion et référencement des e-journaux à l'ULg : historique, état des lieux & perspectives'. It includes fields for 'Document type', 'Discipline(s)', and 'To cite this reference:'. There's a 'Send an email to Renaville, François' button and a form for sending an email. Below that, there are social media sharing options (Facebook, Twitter, LinkedIn, etc.) and a 'Send' button. At the bottom, there's a table for 'File(s) associated to this reference' with columns for File, Version, Size, and Access.

Presentazione del contenuto dell'oggetto rilevato dopo l'accettazione della Licenza d'uso:

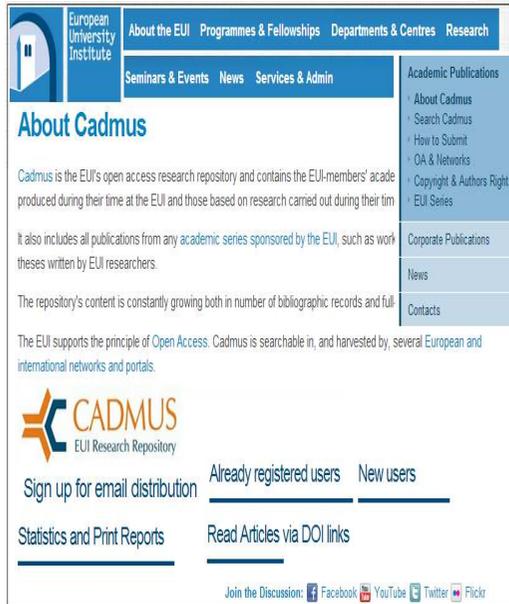
The screenshot shows the 'ORBi User License' text. It starts with 'Thanks for reading and accepting the following us purposes, to enhance teaching or for scientific re'. There are two buttons: 'I accept the license' and 'I do not accept the license'. Below that, there's a section titled 'Ancient Bidialectalism And Bilingualism' with text about the Hellenic world and linguistic diversity. Another section is titled 'The Archaic and Classical Period' with text about evidence of contact with other languages in the Iliad.

V. Comunicazione mediata dal Repository in collegamento con altri sistemi, moduli:

1. A partire dal Repository vi è un'interazione tra suoi contenuti coi contenuti digitali di:
 - a) altri Repository del Belgio,
 - b) aggregatori di contenuti di ricerca come Driver, NARCIS.
2. Modulo/i di E-Mail di autore/i nella scheda descrittiva dell'oggetto.
3. Modulo "Commentary" per i commenti critici e la valutazione dei contenuti nella scheda descrittiva.
4. Possibilità di condividere le informazioni descrittive dell'oggetto depositato attraverso diversi Social Networks.
5. Funzionalità di Feedback per la comunicazione attiva al servizio dell'utente.

¹ La parola latina "orbi" significa "per il mondo"

I. Presentazione dell'Archivio/Repository:



The screenshot shows the Cadmus website interface. At the top, there are navigation tabs: 'About the EUI', 'Programmes & Fellowships', 'Departments & Centres', and 'Research'. Below this, there are sub-tabs for 'Seminars & Events', 'News', and 'Services & Admin'. The main content area features the 'About Cadmus' section, which describes the repository's purpose and search capabilities. There are also links for 'Sign up for email distribution', 'Statistics and Print Reports', and 'Read Articles via DOI links'. Social media icons for Facebook, YouTube, and Twitter are visible at the bottom.

Software: DSpace Open Source
Lingua di presentazione: Inglese

Finalità: raccogliere e preservare il materiale di ricerca digitale/digitalizzato dell'Università di Cadmus. Disseminare la conoscenza sul medesimo materiale tramite l'Internet

Livelli d'accesso:

1. Un accesso autorizzato allo staff accademico e ricercatori dell'Ateneo per il deposito dei propri lavori scientifici

2. Un accesso libero per la ricerca e la consultazione di oggetti (*full-text* e/o *Abstract*) depositati e schede di metadati. È possibile scaricare *full-text* se tale non è impedito dal *Copyright* e/o dal periodo di embargo.

POLICY del Repository + Dichiarazioni sull'OPEN ACCESS + Politiche di editori, COPYRIGHT + Dichiarazioni sui METADATI¹ + Dichiarazioni su FORMATI + trovato + non trovato -

Contenuti di ricerca:

1. Oggetti depositati (13.200): articoli, e-books, rapporti tecnici, tesi, *working papers*, *Courses of the Academy of European Law*.

2. Schede descrittive (record di metadati) per la rappresentazione di oggetti depositati.

II. Aspetti tecnici per la

presentazione di oggetti depositati:

1. Corporate publications: presentazione delle risorse accademiche e informazioni generali sulle attività e programmi dell'EUI raggruppati per anni. Sono sfogliabili online con le funzionalità di "E-Book Viewer" incorporato.

2. Recent submissions: riepilogo regolarmente aggiornato sui nuovi oggetti depositati tramite le funzioni di *Alerting*, *feed RSS/Atom*.

3. Indirizzi Persistenti per ogni oggetto.

4. Statistiche d'uso (accesso, visualizzazione, download) per ogni oggetto/ collezione

III. Ricerca, navigazione:



The screenshot shows the search and browse interface. At the top, there is a 'Search' box with a 'Keyword' field and a 'Search' button. Below this, there is a 'Browse' section with a 'Department of Economics (ECO)' dropdown menu. The 'Browse' section is organized into a grid of categories, including 'Communities & Collections', 'By Issue Date', 'Authors', 'Titles', 'Subjects', and 'Type'. The 'Subjects' category is expanded, showing a list of departments and programs such as 'Department of History and Civilization (HEC)', 'Department of Law (LAW)', 'Department of Political and Social Sciences (SPS)', 'EUI Theses', 'Max Weber Programme (MWP)', 'Robert Schuman Centre for Advanced Studies (RSCAS)', 'Working Papers', and 'Corporate Publications'.

IV. Presentazione dei risultati di ricerca e di navigazione:

Recent Submissions

Issue Date	View	Title	Author(s)	Type of Publication	Series/Report no.	Abstract
2012		From Centre to Province: Changing images of Europe in the writings of Jerzy Stępiński	MIKOLAJEWSKI, Lukasz	Contribution to book		View Abstract
2008		J.-J. Guyonnet de Châteaubourg (1749-1824) ou le commerce des relations	CONCHON, Anne	Book		View Abstract



Presentazione della scheda descrittiva (registrazione attraverso i metadati) dell'oggetto rilevato:

Sotto controllo : *Lecture femminili in Italia nella prima età moderna*



The screenshot shows the descriptive metadata page for a document. It includes fields for 'Title', 'Author', 'Date', 'Publisher', 'Citation', 'Abstract', 'Description', 'URL', and 'ISBN'. The 'Abstract' field contains a detailed summary of the document's content, discussing the role of women in the first modern age. The 'Description' field provides a list of keywords and topics related to the document. At the bottom, there is a 'Files in this item' section showing a file named 'Von_Tippelskirch.JPG' with a size of 43.80Kb and a format of JPEG image.

Presentazione dei contenuti degli oggetti rilevati:



The screenshot shows a collage of document covers and titles. The central focus is 'The Lautsi Papers: Multidisciplinary Reflections on Religious Symbols in the Public School Classroom'. Other visible titles include 'WORKING PAPERS', 'TOLERANCE, PLURALISM AND SOCIAL COHESION: RESPONDING TO THE CHALLENGES OF THE 21ST CENTURY IN EUROPE', 'ACCEPT PLURALISM', and 'MARTINUS NIJHOFF PUBLISHERS'. The collage also features logos for 'Creative Technology Sites' and 'Efficiency Analysis for Panel and Clustered'.

V. Comunicazione mediata dal Repository in collegamento con altri sistemi, moduli:

1. Interazione tra contenuti del *Repository* con contenuti di Archivi aggregatori Driver, DART, NEREUS, PLEADI.

2. Funzionalità di Forum e Feedback per la comunicazione attiva al servizio dell'utente.

¹ <<http://orbi.ulg.ac.be/project?id=103>>



UPCommons: Open Access Portal to the Knowledge of Universitat Politècnica de Catalunya (Spagna):

<<http://upcommons.upc.edu/?locale=en>>

I. Presentazione dell'Archivio/Repository:

The screenshot shows the UPCommons website interface. At the top, there are navigation links for 'Open repositories of UPC' and 'How to take part'. Below this, there are several sections: 'UPCommons' with a description of the portal, 'Access to all contents' with a search bar, and 'Open repositories of UPC' with a list of various digital collections like 'National's Architectural Heritage Archive (AHC)', 'Digital Archive of the School of Architecture of UPC', and 'UPCommons'. There are also buttons for 'Recent Submits' and 'Feedback'.

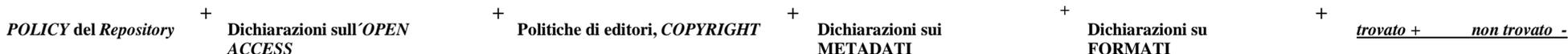
Software: DSpace Open Source
Lingue di presentazione: Inglese, Spagnolo, Catalano

Finalità: funzione di un meta-Archivio che raccoglie, gestisce, preserva e diffonde in rete - a partire da un unico punto d'accesso- la conoscenza sul materiale di ricerca e didattico depositato in tre Archivi digitali dell' *Universitat Politècnica de Catalunya*:

- 1) "Docència", 2) "Recerca", 3) "Altres materials i col·leccions especials"

Livelli d'accesso:

1. Un **accesso autorizzato** per lo staff accademico e ricercatori dell'Ateneo per il deposito dei propri lavori scientifici
2. Un **accesso libero** per la ricerca e la consultazione di oggetti (*full-text* e/o *Abstract*) depositati e schede di metadati. È possibile scaricare *full-text* se tale non è impedito dal *Copyright* e/o dal periodo di embargo.



Contenuti di ricerca:

1. **Oggetti depositati** (45.000): articoli, *e-books* *OpenCourseware*, materiale didattico, tesi di dottorato multimedia (video di lezioni accademiche, conferenze) fotografie, materiale grafico sull'architettura di edifici pubblici di Catalonia, *reports*, *working papers*, *conference proceedings*, presentazioni
2. **Schede descrittive** (record di metadati) per la rappresentazione di oggetti depositati.

II. Aspetti tecnici per la presentazione di oggetti depositati:

1. Caricamento dei riferimenti sull'oggetto nella **scheda personale/CV** dell'autore, a cui si accede dalla scheda descrittiva dell'oggetto depositato
2. Riepilogo regolarmente aggiornato sui **"Recent submissions"** tramite le funzioni di **Alerting, feed RSS/Atom**.
3. **Indirizzi Persistenti** per ogni oggetto
4. **Statistiche d'uso** (accesso, visualizzazione, download) generali e per ogni oggetto.
5. **Importazione delle citazioni** dal servizio SCOPUS
6. **Allegati** al file depositato

III. Navigazione e ricerca dei contenuti:

The screenshot shows the search interface. On the left, there is a 'List of contents' section with a search bar and a 'Recent Submits' button. Below this, there are several categories of documents: 'OpenCourseworks [191 documents]', 'Academic works UPC [14272 documents]', 'E-prints UPC [13798 documents]', and 'Revistes i Congressos UPC [10080 documents]'. On the right, there is a 'Search for' section with a dropdown menu showing 'Dipòsits oberts de la UPC' and a list of search criteria: 'Altres materials i col·leccions especials', 'Videoteca UPC', 'Docència', 'Materials docents UPC - UPCommonsCourseware', 'Treballs acadèmics UPC', 'Recerca', 'E-prints UPC', 'Revistes i congressos UPC', and 'Tesis doctorals - TDX'.

L'impiego delle Classificazioni: *Library of Congress Subject Headings (LCSH)*, *UPC Subject area* per la navigazione a faccette (o multidimensionale).

IV. Presentazione dei risultati di ricerca e di navigazione:

The screenshot shows a search result page. At the top, there are tabs for 'Title', 'Author(s)', 'Contributors', and 'Type'. Below this, there is a list of search results. The first result is 'The stable-abducible argumentation semantics' by Osorio, Miquel; Nieves Sánchez, Juan Carlos; Carballido, José Luis. The second result is 'Stable semantics and inconsistency' by Osorio, Miquel; Zapata, Claudio; Nieves Sánchez, Juan Carlos; Carballido, José Luis. The page also shows the document type as 'Conference report' and the language as 'Catalan'.

Reindirizzamento dall'oggetto selezionato all'Archivio di deposito:

The screenshot shows the TDX website. At the top, there is a search bar and a 'Reindirizza a TDX' button. Below this, there are statistics: '14,428 theses in deposit' and '46,572 searchable theses'. There is also a list of participants from various universities, including 'Universitat de Barcelona', 'Universitat Autònoma de Barcelona', 'Universitat Politècnica de Catalunya', 'Universitat Pompeu Fabra', 'Universitat de Lleida', 'Universitat de València', 'Universitat de Girona', 'Universitat de Vic', 'Universitat Internacional de Catalunya', 'Universitat de les Illes Balears', 'Universitat de Màlaga', 'Universitat de Cantabria', and 'Universitat de Duizhou'.

Presentazione della scheda descrittiva (registrazione attraverso i metadati) dell'oggetto rilevato:

The screenshot shows a detailed document record page. The title is 'The stable-abducible argumentation semantics'. The authors are 'Osorio, Miquel; Nieves Sánchez, Juan Carlos; Carballido, José Luis'. The subjects are 'Argumentation framework', 'Argumentation semantics', 'Logic programming semantics', and 'Programación lógica'. The type is 'Conference report'. The description is: 'We look at a general way of inducing semantics in argumentation theory by means of a mapping models of a normal program, which is constructed in terms of the argumentation framework. In its argumentation semantics called stable abducible which lies in between the stable and the prefer non semantics is that it is nonempty for any argumentation framework, and coincides with the stable whenever this is non-empty. We study some of the properties of this semantic.' There is also a table showing the document's availability in the repository across different countries and dates.

Presentazione del contenuto dell'oggetto rilevato:

The screenshot shows a video player interface. At the top, there is a 'CAPITULO-1' section. Below this, there is a table of contents with the following entries: '1 INTRODUCCION.....', '1.1 INTRODUCCION.....', '1.1.1 Tipologia de la fabrica a considerar en el estudio.....', '1.1.1.1 Material que la constituye: ladrillo, hormigón, piedra.....', and '1.1.1.2 Tipo de unión entre las piezas de la fabrica: con juntas o en seco (sin juntas de mortero).....'. The video player shows a play button and a progress bar.

V. Comunicazione mediata dal Repository in collegamento con altri sistemi, moduli:

1. Interazione tra contenuti di diversi *Repository* accessibili da un'unica interfaccia *UPCommons*, tramite la Classificazione impiegata
2. Interazione tra contenuti di ricerca dell'*UPCommons* con contenuti di:
 - a) cataloghi digitali della *UPC Digital Library*,
 - b) dell'*Anagrafe di Ricerca dell'Ateneo "DRAC"*.
3. Interazione dei contenuti del *Repository "Tesi e Dissertazioni online"* - gestito dal Consorzio *"University Libraries of Catalonia (CBUC)"* con:
 - a) aggregatori delle **tesi del dottorato a livello europeo DART**,
 - b) aggregatori delle **tesi del dottorato a livello internazionale "Chamo Discovery VTLS"**
4. Possibilità di condividere le informazioni sui contenuti depositati attraverso i vari *Social Networks*, a partire dalla scheda descrittiva.
5. Funzionalità di **Feedback** per la comunicazione attiva al servizio dell'utente.



Minho's institutional Repository: RepositoriUM (Portogallo): < <http://repositorium.sdum.uminho.pt/>>

I. Presentazione dell'Archivio/Repository:

Software: OAI Extended Addon DSpace 1.8.2¹
Lingue di presentazione: Inglese, Francese, Spagnolo, Portogallo

Finalità: raccogliere e preservare il materiale di ricerca digitale/digitalizzato dell'Università di Minho. Disseminare la conoscenza sul medesimo materiale tramite l'Internet

Livelli d'accesso:

1. **Un accesso autorizzato** allo staff accademico e ricercatori dell'Ateneo per il deposito dei propri lavori scientifici
2. **Un accesso libero** per la ricerca e la consultazione di oggetti (*full-text* e/o *Abstract*) depositati e schede di metadati. È possibile scaricare *full-text* se tale non è impedito dal *Copyright* e/o dal periodo di embargo.

Contenuti di ricerca:

1. **Oggetti depositati** (20.000): articoli, *working papers*, tesi, *e-book*, *conference proceedings*, materiale didattico, patenti, report, dati di ricerca.
2. **Schede descrittive** (record di metadati) per la rappresentazione di oggetti depositati.

II. Aspetti tecnici per la presentazione di oggetti depositati:

1. I lavori degli autori depositanti sono automaticamente caricati nei loro **CV personali** creati dal servizio incorporato **DeGóis**:

2. **Resent submissions:** riepilogo regolarmente aggiornato sui nuovi oggetti depositati tramite le funzioni di *Alerting*, *feed RSS/Atom*.
3. **Indirizzi Persistenti** per ogni oggetto.
4. **Statistiche d'uso** (accesso, visualizzazione, e *download*) per ogni oggetto.

III. Ricerca, navigazione:

IV. Presentazione dei risultati di ricerca e di navigazione:

Issue Date	Title	Author(s)	Type
2012	School achievement by ethnically and linguistically diverse development programs in Portugal	Moreira, Maria Alfredo	conferenceObject
Sep-2009	Planeamento da urbanização e do turismo em retrospetiva	Sousa, A, Lourenço, Júlia	article
Aug-2011	Pro-drop and theories of pro in the minimalist program and the pronominal-agr hypothesis	Barbosa, Pilar	article

Presentazione della scheda descrittiva (registrazione attraverso i metadati) dell'oggetto rilevato:

Presentazione del contenuto dell'oggetto rilevato:

Livros didáticos de Língua Portuguesa para o Ensino Fundamental: construindo relações entre professores e gêneros discursivos

Adriana Fischer*

Resumo:

O objetivo deste trabalho é o de investigar como e a partir de quais conhecimentos teórico-práticos alunos/professores de um curso de Especialização analisam livros didáticos do Ensino Fundamental quanto à utilização de gêneros discursivos como instrumentos de ensino-aprendizagem em Língua Portuguesa. O instrumento de coleta foi o trabalho final de uma disciplina, ocorrida em julho de 2004, o qual traz as análises de livros didáticos realizadas por eles. Os resultados revelam que os alunos/professores utilizaram, nas análises, alguns aspectos da natureza e constituição dos gêneros. Prevaleceu o enfoque à temática dos textos e às propriedades formais.

Palavras-chave: Livros didáticos. Linguística. Língua portuguesa-Gênero.

* Doutoranda (2004/1-2008/1) do Programa de Pós-graduação em Letras/Linguística da UFSC. Professora no Centro Universitário de Brusque (UNIFEBE) e Universidade para o Desenvolvimento do Alto Vale do Itajaí (UNIDAVI) – Rio do Sul-SC. Bolsista CAPES (2004-2008).

PERSPECTIVA, Florianópolis, v. 24, n. 2, p. 533-567, jul./dez. 2006

<http://www.perspectiva.ufsc.br>

V. Comunicazione mediata dal Repository in collegamento con altri sistemi, moduli:

1. Interazione tra contenuti del *Repository* con contenuti di:
 - a) Catalogo della Biblioteca universitaria;
 - b) Portale RCAAAP (aggregatore di 47 Archivi *Open Access* del Portogallo);
 - c) dell'infrastruttura OpenAIRE (collezioni di ricerca del FP7);
2. possibilità di condividere le informazioni sull'oggetto depositato tramite vari *Social Networks*;
3. Funzionalità di *Feedback* per la comunicazione attiva al servizio dell'utente.

POLICY del Repository + Dichiarazioni sull'OPEN ACCESS + Politiche di editori, COPYRIGHT + Dichiarazioni su METADATI + Dichiarazioni su FORMATI - trovato + non trovato -

¹ Il Software è sviluppato in conformità con i requisiti delle Linee Guida OpenAIRE (*Open Access Infrastructure for Research in Europe*). Metadati: <<http://orbi.ulg.ac.be/project?id=103>>



Archive Ouverte HAL-INRIA (INRIA a CCSD electronic Archive server) (Francia): <<http://hal.inria.fr/>>

I. Presentazione dell'Archivio/Repository:

The screenshot shows the HAL-INRIA website interface. At the top, there's a navigation bar with 'Déposer', 'Consulter', 'Rechercher', 'Services', and 'Aide'. Below it, there's a search bar and a 'Zoom sur...' section. A 'RANKING WEB OF REPOSITORIES' section is visible, along with a 'Consultez les derniers dépôts' section. The interface is in French and English.

Software: HAL
Lingue di presentazione: Francese, Inglese

Finalità: funzione di un meta-Archivio che - a partire da un unico punto d'accesso - raccoglie, gestisce e diffonde in rete la conoscenza sul materiale scientifico delle istituzioni di istruzione e di ricerca francesi che hanno firmato a favore del protocollo HAL per aderire alla piattaforma comune di ricerca francese disponibile in *Open Access*.

Livelli d'accesso:

1. Un **accesso autorizzato** allo staff accademico, ai ricercatori e agli studenti per il deposito del materiale.
2. Un **accesso libero** per la ricerca e la consultazione di oggetti (*full-text* e/o *Abstract*) depositati e schede di metadati. È possibile scaricare *full-text* se tale non è impedito dal *Copyright* e/o dal periodo di embargo.

Contenuti di ricerca:

1. **Oggetti depositati** (240.000): articoli, *e-books*, rapporti tecnici, interventi in convegni e conferenze, tesi di master e di dottorato, *working papers*, materiale didattico, curatele, multimedia
2. **Schede descrittive** (record di metadati) per la rappresentazione di oggetti depositati.

II. Aspetti tecnici per la presentazione di oggetti depositati:

1. **“Latest documents deposited”:** riepilogo regolarmente aggiornato sui nuovi oggetti depositati tramite le funzioni di *Alerting*, *feed RSS/Atom*.
2. **Linking** degli **oggetti rilevati** che riportano *Homepage* e *CV* degli autori

The screenshot shows a user profile for Salma Lahbabi, PhD student at CNRS, Université de Cergy. It features navigation tabs for 'Curriculum vitae', 'Publications and preprints', and 'Teaching'. There are also links for 'Contact information' and 'Personal webpage'.

3. **Oggetti affiancati da file allegati**
4. **Indirizzi Persistenti** per ogni oggetto depositato.

III. Navigazione e ricerca dei contenuti:

The screenshot shows the HAL search interface with various filters and navigation options. It includes sections for 'Browse > News (the last 30) ..', 'Browse > by domain ..', 'Browse > by collection ..', and 'Browse > by writing date'. There are also links for 'fr.arXiv.org' and 'Quantitative Finance and Statistics'.

IV. Presentazione dei risultati di ricerca e di navigazione:

The screenshot shows a search results page for 'linguistics'. It displays a list of documents, including 'A Generative Montogavian Lexicon for Polyssemous Deverbal Nouns' and 'Den sociale konstruktion af idéer: en multimodal interaktionsanalytisk undersøgelse af hvordan idéer skab differentierende eller affilerende responsformat'. There are also links for 'See short view' and 'Export'.

Presentazione della scheda descrittiva (registrazione attraverso i metadati) dell'oggetto rilevato:

The screenshot shows a document detail page for 'Letter counting: a stem cell for Cryptology, Quantitative Linguistics, and Statistics'. It includes a title, author information (Bernard Ycart), a description, and an attached file list with details like 'hal-00758644, version 1' and 'LF1.pdf (192.3 KB)'. There are also links for 'Associated documents' and 'Bibliographic reference'.

Presentazione del contenuto dell'oggetto rilevato:

The screenshot shows the abstract and introduction of the document 'Partitioning, Ordering, and Load Balancing in a Hierarchically Parallel Hybrid Linear Solver'. The abstract discusses the challenges of solving large linear systems and the proposed hierarchical parallel solver. The introduction provides more details on the solver's architecture and performance.

V. Comunicazione mediata dal Repository in collegamento con altri sistemi, moduli:

1. Dalla scheda descrittiva è possibile risalire ai contenuti correlati dell'Archivio globale *arXiv* del materiale di ricerca in discipline tecniche.
2. È possibile esportare le informazioni sull'oggetto in *arXiv* e *RefWorks* :

The screenshot shows the 'Documents associés' section with links to 'arXiv: 1211.6847' and 'RefWorks'. There are also links for 'Bibtex', 'EndNote', 'TEI', and 'RefWorks'.

POLICY del Repository + Dichiarazioni sull'OPEN ACCESS + Politiche di editori, COPYRIGHT + Dichiarazioni sui METADATI + Dichiarazioni su FORMATI + trovato + non trovato -



Repository of the University of Southampton ePrints Soton (Inghilterra): <<http://eprints.soton.ac.uk/>>

I. Presentazione dell'Archivio/Repository:

Software: [EPrints Open Source](#)

Lingua di presentazione: Inglese

Finalità: raccogliere e disseminare tramite l'Internet il materiale di ricerca digitale dell'Università di Southampton.

Livelli d'accesso:

1. Un accesso autorizzato per lo staff accademico e ricercatori dell'Ateneo per il deposito dei propri lavori scientifici

2. Un accesso libero per la ricerca e la consultazione di oggetti (*full-text* e/o *Abstract*) depositati e schede di metadati. È possibile scaricare *full-text* se tale non è impedito dal *Copyright* e/o dal periodo di embargo

Contenuti di ricerca:

1. **Oggetti depositati** (35.000): articoli, *e-books* e/o loro capitoli, rapporti tecnici, interventi in convegni e conferenze, tesi di master e di dottorato, *working papers*, multimedia
2. **Schede descrittive** (record di metadati) per la rappresentazione di oggetti depositati.

II. Aspetti tecnici a supporto di presentazione di oggetti depositati:

1. **"Latest additions"**: un riepilogo regolarmente aggiornato sui nuovi oggetti depositati tramite le funzioni di *Alerting*, *feed RSS/Atom*.
2. La tecnologia *Linked Data* a supporto delle connessioni tra l'identificatore dell'oggetto depositato, il suo **tema** (soggetto) e la **divisione** dell'Università della provenienza dell'oggetto.

III. Navigazione e ricerca dei contenuti:

L'impiego della Classificazione *"Library of Congress Subject Areas"* per la navigazione a faccette (o multidimensionale):

POLICY del Repository

Dichiarazioni sull'OPEN ACCESS

Politiche di editori, COPYRIGHT

IV. Presentazione dei risultati di ricerca e di navigazione:

Presentazione della scheda descrittiva (registrazione attraverso i metadati) dell'oggetto rilevato:

Dichiarazioni su METADATI

Dichiarazioni su FORMATI

Presentazione del contenuto dell'oggetto rilevato:

V. Comunicazione mediata dal Repository in collegamento con altri sistemi, moduli:

1. L'interazione diretta dei contenuti depositati con i contenuti della piattaforma web della Biblioteca dell'Ateneo

2. A partire dalla scheda descrittiva è possibile:

a) condividere le informazioni sull'oggetto depositato tramite diversi *Social Networks*;

b) accedere ai contenuti correlati in *Google Scholar*

trovato + non trovato -



I. Presentazione dell'Archivio/Repository:

The screenshot shows the UCL Discovery website interface. At the top, there is a search bar with 'Everything' selected and 'Author' as a filter. Below the search bar, there are navigation menus for 'Browse by' (Department, Year, Latest) and 'For everyone' (About UCL Discovery, Open Access, Using UCL Discovery). A central section titled 'UCL Research' features a video player and text about UCL's research. On the right, there are statistics for top downloads and a 'Featured' section with a link to 'Multivariate Money'. At the bottom, there is a 'For UCL authors' section with links to maintain publications and deposit policies.

Software: EPrints *Open Source*
Lingua di presentazione: Inglese

Finalità: raccogliere e preservare il materiale di ricerca digitale dell'Ateneo. Disseminare la conoscenza sul medesimo materiale tramite l'Internet

Livelli d'accesso:

1. Un **accesso autorizzato** per: **a)** per lo staff accademico e ricercatori dell'Ateneo per il deposito dei propri lavori scientifici; **b)** la consultazione di *full-text* per alcuni oggetti.
2. Un **accesso libero** per la ricerca e la consultazione dei contenuti di ricerca. È possibile scaricare *full-text* disponibile in modalità *Open Access*

Contenuti di ricerca:

1. **Oggetti depositati** (270.000): articoli presentati a conferenze o pubblicati su rivista, *working papers*, *e-books*, rapporti di ricerca, tesi di dottorato, multimedia
2. **Schede descrittive** (record di metadati) per la rappresentazione di oggetti depositati.

II. Aspetti tecnici per la presentazione di oggetti depositati:

1. **“Latest additions”**: riepilogo regolarmente aggiornato sui nuovi oggetti depositati tramite le funzioni di *Alerting*, *feed RSS/Atom*.
2. **Statistiche d'uso** (accesso, visualizzazione, download)
3. Esportazione di riferimenti di oggetti depositati su **“Personal publication lists”**

III. Navigazione e ricerca dei contenuti:

The screenshot shows the search results page for 'Mamas, T (2012) Investigation College London'. It includes a search bar, a 'Detailed search' button, and a list of results. The first result is 'Mamas, T (2012) Investigation College London'. Below the search results, there is a 'Browse by UCL Departments and Centres' section with a tree view showing various departments like 'School of Arts and Social Sciences (04182)', 'Faculty of Arts and Humanities (10919)', 'Faculty of Law (0446)', 'Faculty of Social and Historical Sciences (18057)', and 'SEERC (School of Slavonic and East European Studies) (2727)'. A 'Yesterday' section shows a result for 'Mamas, T (2012) Investigation College London'.

IV. Presentazione dei risultati di ricerca e di navigazione:

The screenshot shows a search result for 'The subjunctive conundrum in English' by Aerts, B (2012). The result includes the title, author, and a link to the full text. A red box highlights the text 'An open access version is available from UCL Discovery'.

Presentazione della scheda descrittiva (registrazione attraverso i metadati) dell'oggetto rilevato:

The screenshot shows the detailed description page for 'Renaissance Queenship in William Shakespeare's English History Plays' by Chang, Y.-C. (2012). The page includes the title, author, and a link to the full text. A red box highlights the text 'An open access version is available from UCL'. The page also includes an abstract and a list of keywords.

Presentazione del contenuto dell'oggetto rilevato:

The screenshot shows the content page for 'Grammatici certant' by Rodney Huddleston & Geoffrey K. Pullum. The page includes the title, authors, and a link to the full text. A red box highlights the text 'An open access version is available from UCL Discovery'.

V. Comunicazione mediata dal Repository in collegamento con altri sistemi, moduli:

1. Servizio di *Blog* per la comunicazione attiva al servizio dell'utente:

The screenshot shows the UCL Discovery Blog page. The main content is a notice about a scheduled outage for the Research Publications Service (RPS) on Tuesday 26th February 2013 from 15:00. The page also includes a sidebar with a 'UCL Discovery blog email alerts' section and a 'Subscribe' button.

2. Possibilità di condividere le informazioni sul *Repository* tramite i vari *Social Networks*

POLICY del Repository

Dichiarazioni sull'OPEN ACCESS

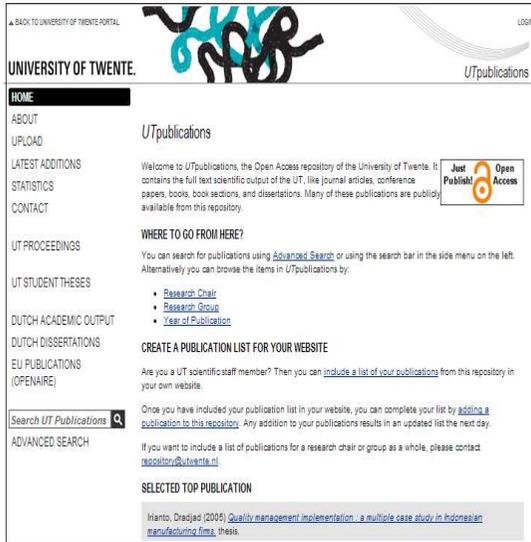
Politiche di editori, COPYRIGHT

Dichiarazioni su METADATI

Dichiarazioni su FORMATI

trovato + non trovato -

I. Presentazione dell'Archivio/Repository:



The screenshot shows the homepage of the University of Twente Publications repository. It features a navigation menu on the left with options like HOME, ABOUT, UPLOAD, LATEST ADDITIONS, STATISTICS, CONTACT, UT PROCEEDINGS, UT STUDENT THESES, DUTCH ACADEMIC OUTPUT, DUTCH DISSERTATIONS, EU PUBLICATIONS (OPENAIRE), and ADVANCED SEARCH. The main content area includes a welcome message, a search bar, and a section for 'WHERE TO GO FROM HERE?' with links to Research Chair, Research Group, and Year of Publication. There is also a section for 'CREATE A PUBLICATION LIST FOR YOUR WEBSITE' and a 'SELECTED TOP PUBLICATION' section.

Software: EPrints *Open Source*

Lingua di presentazione: Inglese

Finalità: raccogliere e preservare il materiale di ricerca digitale dell'Università di *Twente*. Disseminare la conoscenza sul medesimo materiale tramite l'Internet

Livelli d'accesso:

1. Un accesso autorizzato per lo staff accademico e ricercatori dell'Ateneo per il deposito dei propri lavori scientifici

2. Un accesso libero per la ricerca e la consultazione di oggetti (*full-text* e/o *Abstract*) depositati e schede di metadati. È possibile scaricare *full-text* se tale non è impedito dal *Copyright* e/o dal periodo di embargo.

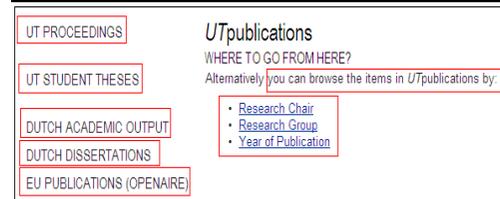
Contenuti di ricerca:

- 1. Oggetti depositati (35.000):** articoli, *e-books*, rapporti tecnici, interventi in convegni e conferenze, tesi di master e di dottorato, *working papers*.
- 2. Schede descrittive** (record di metadati) per la rappresentazione di oggetti depositati.

II. Aspetti tecnici per la presentazione di oggetti depositati:

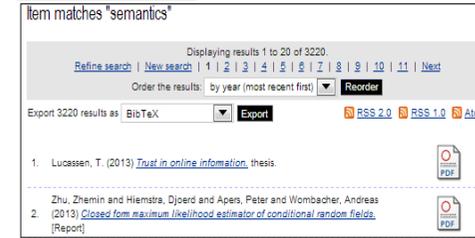
- 1. "Latest additions":** riepilogo regolarmente aggiornato sui nuovi oggetti depositati tramite le funzioni di *Alerting*, *feed RSS/Atom*.
 - Collegamento dell'oggetto depositato con uno o più **Allegati**.
 - 3. Statistiche d'uso** (accesso, visualizzazione, *download*) generali e per ogni oggetto depositato.
 - Creazione automatica delle **liste di pubblicazioni** con il loro **export** su **website personali** di autori depositanti
- CREATE A PUBLICATION LIST FOR YOUR WEBSITE
- Are you a UT scientific staff member? Then you can include a list of your publications from this repository in your own website.
- Once you have included your publication list in your website, you can complete your list by adding a publication to this repository. Any addition to your publications results in an updated list the next day.
- 5. Indirizzi Persistenti** per ogni oggetto depositato e la conservazione di oggetti a lungo termine presso la *National Library of the Netherlands*.

III. Navigazione e ricerca dei contenuti



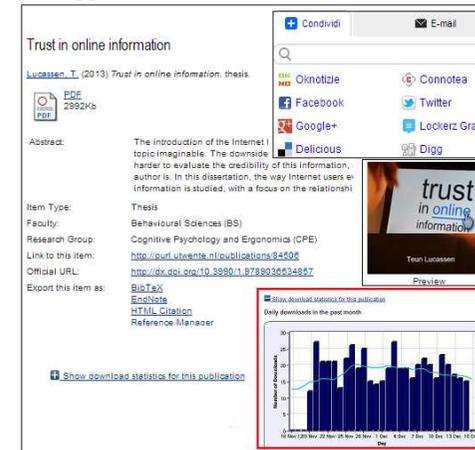
The screenshot shows the navigation menu of the UTpublications website. It includes links for UT PROCEEDINGS, UT STUDENT THESES, DUTCH ACADEMIC OUTPUT, DUTCH DISSERTATIONS, and EU PUBLICATIONS (OPENAIRE). There is also a section for 'WHERE TO GO FROM HERE?' with links to Research Chair, Research Group, and Year of Publication.

IV. Presentazione dei risultati di ricerca e di navigazione:



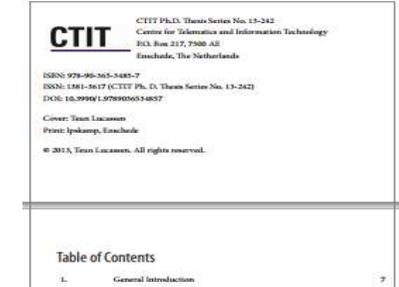
The screenshot shows search results for the term 'semantics'. It displays 3220 results, with the first two results listed: 1. Lucassen, T. (2013) *Trust in online information*, thesis. 2. (2013) *Closed form maximum likelihood estimator of conditional random fields*. [Report]. There are options to export results as BibTeX, RSS 2.0, RSS 1.0, and Atom.

Presentazione della scheda descrittiva (registrazione attraverso i metadati) dell'oggetto rilevato:



The screenshot shows the descriptive page for the thesis 'Trust in online information' by Lucassen, T. (2013). It includes the abstract, item type (Thesis), faculty (Behavioural Sciences (BS)), research group (Cognitive Psychology and Ergonomics (CPE)), link to the item, official URL, and export options (BibTeX, EPrints, HTML Citation, Reference Manager). There is also a section for 'Show download statistics for this publication' with a bar chart showing daily downloads in the past month.

Presentazione del contenuto dell'oggetto rilevato:



The screenshot shows the table of contents for the thesis 'Trust in online information'. It includes the title, author (T. Lucassen), and a list of contents: L. General Introduction.

V. Comunicazione mediata dal Repository in collegamento con altri sistemi, moduli:

- L'interazione diretta dei contenuti depositati con i contenuti degli aggregatori del materiale di ricerca come *NARCIS_Dutch Academic Output*; *NARCIS_Dutch Dissertations*; *OpenAIRE*:



The screenshot shows the NARCIS and OpenAIRE search interfaces. NARCIS is a gateway to scholarly information in the Netherlands, and OpenAIRE is an open access infrastructure for research in Europe. Both interfaces show search bars and navigation options.

- A partire dalla scheda descrittiva è possibile: **a)** accedere ai contenuti correlati in *Google Books search*, *Google Scholar*; **b)** condividere contenuti tramite vari *Social Networks* e *e-mail*



Munich Personal RePEc Archive. Ludwig-Maximilians-Universität München (Germania):

<http://mpra.ub.uni-muenchen.de/>

Presentazione dell'Archivio/Repository:

The screenshot shows the MPRA (Munich Personal RePEc Archive) website. It features a navigation bar with links for Home, Browse, Search, FAQ, About, and Help. A search bar is prominently displayed. Below the search bar, there are sections for 'BROWSE' (By Author, By Institution, By Year, etc.) and 'SERVICES' (RSS-Feed, etc.). A 'WELCOME!' message is visible, along with information about the repository's purpose and submission guidelines.

Software: EPrints *Open Source* 3.0.5¹

Lingua di presentazione: Inglese

Finalità: fornire uno tra i servizi della più estesa rete delle collezioni digitali di lavori scientifici in economia *RePEc (Research Papers in Economics)*², un progetto a base volontaria ispirato al movimento per il Software *Open Source* a cui sono aderiti più di **1400 Archivi** da **75 paesi**, i quali entro il 2013 hanno condiviso tra loro circa **1, 5 milioni** oggetti digitali

Livelli d'accesso:

1. Un accesso autorizzato per i ricercatori affinché possano disseminare i propri lavori scientifici tramite l'Archivio nella rete RePEc.

2. Un accesso libero per la ricerca e la consultazione di oggetti (*full-text* e/o *Abstract*) depositati e schede di metadati. È possibile scaricare *full-text* se tale non è impedito dal *Copyright* e/o dal periodo di embargo.

Contenuti di ricerca:

- 1. Oggetti depositati (25.000):** *working papers, discussion papers, e-books*, rapporti di ricerca, articoli presentati a conferenze o pubblicati su qualsiasi rivista
- 2. Schede descrittive** (record di metadati) per la rappresentazione di oggetti depositati.

II. Aspetti tecnici per la presentazione di oggetti depositati:

1. Il materiale di ricerca degli autori registrati nell'Archivio aggregatore RePEc³ sarà automaticamente caricato nei loro "CV RePEc" personali

2. Resent submissions: riepilogo regolarmente aggiornato sui nuovi oggetti depositati tramite le funzioni di *Alerting, feed RSS/Atom*.

3. Statistiche d'uso (accesso e *download*) per ogni oggetto

III. Navigazione e ricerca dei contenuti:

L'impiego della *JEL Classification* per la navigazione a faccette (o multidimensionale):

The screenshot shows a hierarchical tree of the JEL Classification. The root is 'A - General Economics and Teaching (1226)'. Under it, there is 'A2 - Economics Education and Teaching of Economics (288)'. Further sub-categories include 'A20 - General (95)', 'A21 - Pre-college (18)', 'A22 - Undergraduate (56)', 'A23 - Graduate (68)', and 'A29 - Other (16)'. The 'Browse' button is highlighted at the top.

IV. Presentazione dei risultati di ricerca e di navigazione:

The screenshot shows a search results page. At the top, there are filters for 'JEL Classification (21940)', 'History of Economic Thought, Methodology, and Historical Approaches (1137)', 'History of Economic Thought through 1933 (253)', and 'Socialist: Marxist (64)'. Below this, a list of papers is shown, including 'L'ARTICULATION DES MODES DE PRODUCTION CHEZ GEORG ELWERT' by Ahoiounkanzon, Michel (2009) and 'The Variable Time: crucial to understand' by Khumalo, Bhakuzulu (2008).

Presentazione della scheda descrittiva (registrazione attraverso i metadati) dell'oggetto rilevato:

The screenshot shows a detailed record page for the paper 'Learning from physics education research: Lessons for economics education' by Simkins, Scott P. and Maier, Mark H. (2008). It includes a PDF download link, an abstract, and a metadata table. The metadata table lists: Item Type: MPRA Paper; Original Title: Learning from physics education research: Lessons for economics education; Language: English; Keywords: economic education; physics education research (PER); research-based teach instruction; just-in-time teaching; interactive lecture demonstration; Subjects: A - General Economics and Teaching > A2 - Economics Education and Teaching; Item ID: 9314; Depositing User: Scott Simkins; Date Deposited: 19. Jul 2008 08:27; Last Modified: 19. Feb 2013 08:56; References: Barkley, Elizabeth F.; Cross, K. Patricia and Major, Claire Howell. Collaborative

Presentazione del contenuto dell'oggetto rilevato:

The screenshot shows the abstract and keywords for the paper. The title is 'Learning from Physics Education Research: Lessons for Economics Education' by Mark Maier* and Scott Simkins**. The abstract states: 'We believe that economists have much to learn from educational research practices and related pedagogical innovations in other disciplines, in particular physics education. In this paper we identify three key features of physics education research that distinguish it from economics education research - (1) the intentional grounding of physics education research in learning science principles, (2) a shared conceptual research framework focused on how students learn physics concepts, and (3) a cumulative process of knowledge-building in the discipline - and describe their influence on new teaching pedagogies, instructional activities, and curricular design in physics education. In addition, we highlight four specific examples of successful pedagogical innovations drawn from physics education - context-rich problems, concept tests, just-in-time teaching, and interactive lecture demonstrations - and illustrate how these practices can be adapted for economic education.' Keywords include: economic education, physics education research (PER), research-based teaching, just-in-time teaching, interactive lecture demonstration. JEL Classification: A2 - Economic Education and Teaching of Economics. PDF link: MPRA_paper_9314.pdf.

V. Comunicazione mediata dal Repository in collegamento con altri sistemi, moduli:

1. Interazione tra contenuti del *Repository* con contenuti della:
 - a) rete degli Archivi *RePEc*;
 - b) Biblioteca del'Università di *Munich*;
 - c) libreria digitale della *University of California*.

2. Servizio di *Blog* per la comunicazione attiva al servizio dell'utente.

POLICY del Repository + Dichiarazioni sull'OPEN ACCESS + Politiche di editori, COPYRIGHT + Dichiarazioni su METADATI + Dichiarazioni su FORMATI + trovato + non trovato -

¹ <http://www.eprints.org/>

² RePEc (<http://repec.org/>) assicura la maggior circolazione possibile del proprio materiale, un obiettivo che nessun sistema a pagamento o editore privato può fornire. I contenuti di RePEc sono indicizzati dai siti/servizi web come IDEAS (<http://ideas.repec.org/>), EconPapers (<http://econpapers.repec.org/>), NetEc (<http://netec.mcc.ac.uk/NetEc.html>), Inomics (<http://www.inomics.com/>), EconLit (<http://www.aeaweb.org/econlit/>), EconStor (<http://www.econstor.eu/>), Microsoft Academic Search (<http://academic.research.microsoft.com/>), Scirus for scientific information only (<http://www.scirus.com/>), Sciverse (<http://www.hub.sciverse.com/action/home>); NEP (*New Economic Papers*: <http://nep.repec.org/>). RePEc assicura la maggior circolazione possibile del proprio materiale, un obiettivo che nessun sistema a pagamento o editore privato può fornire.

³ <http://authors.repec.org/>



I. Presentazione dell'Archivio/Repository:

The screenshot shows the AIR website interface. On the left, there's a search bar and navigation menu. The main content area features 'OPEN ACCESS' information and a list of categories under 'Scorri' (Browse). A sidebar on the right lists 'I vantaggi per gli autori, le istituzioni e i lettori' and 'Per approfondire: AIR e l'Open Access'.

Software: SURplus (personalizzazione del Software DSpace *Open Source*) – modulo SURplus *Open Archive*
Lingua di presentazione: Italiano, Inglese

Finalità: 1. raccogliere e preservare il materiale di ricerca digitale dell'Ateneo. Disseminare la conoscenza sul medesimo materiale tramite l'Internet;

2. presentare l'**Anagrafe della ricerca** dell'Ateneo¹. I dati dell' AIR:

- a) vengono inviati quotidianamente al sito docente del MIUR;
- b) alimentano le **pagine dei docenti o dei dipartimenti**;
- c) sono utilizzati dall'Amministrazione centrale per la **valutazione delle strutture di ricerca e delle Scuole di Dottorato** e per i processi di **assegnazione dei fondi**;
- d) sono usati per la **redazione delle Relazioni Annuali dei dipartimenti** e per campagne di valutazione promosse da singoli dipartimenti o facoltà;
- e) possono **alimentare i siti web dei dipartimenti o le pagine personali** dei ricercatori;
- f) possono essere esportati dai ricercatori in vari formati, per esempio per l'**inclusione nei propri CV**.

POLICY del Repository + **Dichiarazioni sull'OPEN ACCESS**

Livelli d'accesso:

1. **Un accesso autorizzato** per staff accademico e ricercatori dell'Ateneo per il deposito dei propri lavori scientifici

2. **Un accesso libero** per la ricerca e la consultazione di oggetti (*full-text* e/o *Abstract*) depositati e schede di metadati. È possibile scaricare *full-text* se tale non è impedito dal *Copyright* e/o dal periodo di embargo.

Contenuti di ricerca:

- 1. **Oggetti depositati** (130.000): articoli presentati a conferenze o pubblicati su rivista, *working papers*, *e-books*, rapporti di ricerca, norma o brevetto, mappe, tesi di dottorato, multimedia
- 2. **Schede descrittive** (record di metadati) per la rappresentazione di oggetti depositati.

II. Aspetti tecnici per la presentazione di oggetti depositati:

- 1. **Indirizzi Persistenti** per ogni oggetto
- 2. **Ultime immissioni:** riepilogo regolarmente aggiornato sui nuovi oggetti depositati tramite le funzioni di *Alerting, feed RSS/Atom*
- 3. **Statistiche d'uso** (accesso, visualizzazione, *download*) per ogni oggetto
- 4. **Oggetti affiancati da file allegati**

III. Navigazione e ricerca dei contenuti:

L'impiego della Classificazione dei Settori Scientifico-Disciplinari:

The screenshot shows the 'Scorri' (Browse) menu with a search bar and a list of categories: Autore, Titolo, Titolo di periodico, Settore scientifico disciplinare, and Tipologia. The 'Settore scientifico disciplinare' is expanded to show a list of categories like L-ANT/01 - Preistoria e Protostoria, L-ANT/02 - Storia Greca, etc.

+ **Dichiarazioni Politiche di editori, COPYPRIGHT**

IV. Presentazione dei risultati di ricerca e di navigazione:

The screenshot shows search results. A table lists records with columns for Data, Titolo, Autori, Tipologia, File allegati, and Abstract. Below the table, there's a 'Raffina i risultati' (Refine results) section with filters for 'Autori UniMI', 'Settori disciplinari (italiano)', 'Editori', 'Parole chiave', 'Riviste', 'Tipologie', and 'Full-text'.

Presentazione della scheda descrittiva (registrazione attraverso i metadati) dell'oggetto rilevato:

The screenshot shows a detailed record page for a document. It includes a title, author, and publisher information. There are two bar charts: 'Statistics for item 1842/6087' showing 'No. of Page views' and 'No. of page views per month', and another showing 'No. of Downloads' and 'No. of downloads per month'. The page also mentions 'This item is licensed under a Creative Commons License'.

+ **Dichiarazioni su METADATI** + **Dichiarazioni su FORMATI** + **trovato + non trovato -**

Presentazione del contenuto dell'oggetto rilevato:

The screenshot shows a content page for a 'Threnodic Elegy in Sparta' by Cecilia Nobili. The page features a large 'M' at the start of the text, which discusses the new elegy of Simonides for the fallen at Plataea. It includes a citation: 'A. Aloni, "L'elegia di Simonide dedicata alla battaglia di Plataea," ZPE 102 (1994) 9-22, and "The Poem of Simonides' Plataea Elegy and the Circumstances of its Performance," in D. Boedeker and D. Sider (eds.), The New Simonides: Contexts of Praise and Desire (New York 2001) 86-105; C. O. Greek, Roman, and Byzantine Studies 51 (2011) 26-48 © 2011 Cecilia Nobili'.

V. Comunicazione mediata dal Repository in collegamento con altri sistemi, moduli:

- 1. Attraverso SURplus l' AIR si interfaccia con:
 - a) sistemi gestionali interni dell'Ateneo;
 - b) Archivi bibliografici della Biblioteca dell'Università;
 - c) "Sito docente" del sistema del Ministero MIUR
- 2. Servizio di *Blog* per la comunicazione attiva al servizio dell'utente.
- 3. Possibilità di condividere le informazioni sull'oggetto depositato tramite vari *Social Networks*

¹ Cassella, Maria, Bozzarelli, Oriana. *Open Access, Archivi Istituzionali e valutazione della ricerca*, Torino 24 ottobre, 2011, http://eprints.rclis.org/bitstream/10760/16775/1/OA2011_AIBnotizie.pdf



Publication Open Repository Torino (PORTO). Politecnico di Torino (Italia):

[<http://porto.polito.it/>](http://porto.polito.it/)

I. Presentazione dell'Archivio/Repository:

Software: EPrints *Open Source* collegato con l'applicativo del Politecnico "Cineca UGOV-CATALOGO della RICERCA" da cui avviene il caricamento di oggetti digitali nel *Repository*

Lingue di presentazione: Italiano, Inglese

Finalità: raccogliere e presentare la produzione scientifica del Politecnico, incluso obbligatoriamente quella finanziata dai fondi europei. Disseminare la conoscenza sul medesimo materiale tramite l'Internet

Livelli d'accesso:

1. Un **accesso autorizzato** al personale amministrativo del *Repository* per i controllo dell'inserimento automatico dei dati (adattati alle caratteristiche dei contenuti di un Archivio Aperto) tramite UGOV-CATALOGO della RICERCA"
2. Un **accesso libero** per la ricerca e la consultazione di oggetti (e/o *Abstract*) depositati e delle schede di metadati. È possibile scaricare *full-text* se tale non è impedito dal *Copyright* e/o dal periodo di embargo.

Contenuti di ricerca:

1. **Oggetti depositati** (70.000): articoli e interventi in convegni e conferenze (74%), *e-books*, rapporti tecnici, tesi di dottorato, *working papers*, brevetti, curatele
2. **Schede descrittive** (record di metadati) per la rappresentazione di oggetti depositati.

II. Aspetti tecnici per la presentazione di oggetti depositati:

1. **Oggetti affiancati da file allegati**
2. **"Ultime pubblicazioni ad accesso aperto"**, **"Ultimi titoli inseriti"**; riepilogo regolarmente aggiornato sui nuovi oggetti depositati tramite le funzioni di *Alerting*, *feed RSS/Atom*.
3. **Indirizzi Persistenti** per ogni oggetto depositato.
4. **Statistiche d'uso** (accesso, visualizzazione, *download*) generali.
5. **Top Periodici**

III. Navigazione e ricerca dei contenuti:

L'impiego della Classificazione MIUR dei Settori Scientifico-Disciplinari.

IV. Presentazione dei risultati di ricerca e di navigazione:

Presentazione della scheda descrittiva (registrazione attraverso i metadati) dell'oggetto rilevato:

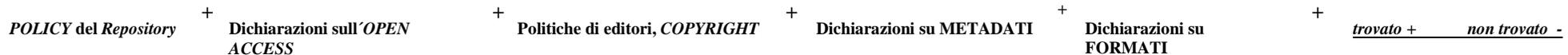
Discussion of the De Generatione Sonorum, a treatise on sound and phonetics by Robert Grosseteste

Presentazione del contenuto dell'oggetto rilevato:

V. Comunicazione mediata dal Repository in collegamento con altri sistemi, moduli:

1. Interazione diretta dei contenuti con "UGOV-CATALOGO della RICERCA" collegato con la Valutazione Interna e Ministeriale: il "Sito docente" del MiUR è alimentato, per le Università che fanno parte del Cineca, solo tramite U-GOV.

2. Possibilità di condividere le informazioni sui contenuti depositati tramite vari *Social Networks*, a partire dalla scheda descrittiva.





MDC (Digital Memory of Catalonia). Archivio fotografico *Open Access* del patrimonio storico di Catalunya (Spagna): <<http://mdc1.cbuc.cat/index.php>>

I. Presentazione dell'Archivio/Repository:

The screenshot shows the MDC website interface. At the top, there are navigation links: Home, What is it?, Collections, Participating institutions, Future collections, and Technical information. Below this, there's a search bar with a 'Search' button and an 'Advanced search' link. The main content area features a collection titled 'AFECTOS RENDIDOS' with a thumbnail image of a woman. To the right, there's a 'Novetats Ara podreu fer' section with a map of Catalonia. At the bottom, there are links for 'Legal note', 'Sitemap', and 'mdc@cbuc.cat'.

Software: CONTENTdm OCLC commercial

The screenshot shows the CONTENTdm software interface. It displays a list of metadata fields for a digital object, including 'Best Practices for CONTENTdm and other OAI-PMH compliant repositories: creating sharable metadata'. The interface includes a search bar, a 'Choose between:' section with options for 'Advanced search inside ARCA journals' and 'Advanced search inside the rest of the collections', and a 'Navegazione di contenuti per:' section with options for 'a) le collezioni delle istituzioni partecipative:' and 'b) keywords (esposti in Tags Cloud) più utilizzati per la ricerca dei contenuti;'. There is also a 'c) collegamenti sviluppati con Linked Data per collegare gli oggetti sulla base dei metadati "Autore" e "Soggetto" e del parametro geografico delle istituzioni:' section.

Lingue di presentazione: Inglese, Spagnolo, Catalano

Finalità: funzione di un meta-Archivio che - a partire da un unico punto d'accesso - raccoglie, gestisce, preserva e diffonde in rete in modalità *Open Access* la conoscenza sul materiale digitalizzato di valore storico della Catalunya, fornito da 24 istituzioni cooperanti (*Consorci de Biblioteques Universitàries de Catalunya, CBUC*)

Livelli d'accesso:

Un accesso autorizzato al personale amministrativo per il deposito dei contenuti nel Repository.

2. Un accesso libero per la ricerca, la consultazione e *download* del materiale depositato

POLICY del Repository

Dichiarazioni sull'OPEN ACCESS

Politiche di editori, COPYRIGHT

IV. Presentazione dei risultati di ricerca e di navigazione:

The screenshot shows search results for 'Fons de la Cambra de Comerç de Barcelona (Universitat Pompeu Fabra)'. It includes a list of collections with titles like 'Fons de la Guerra de Successió (1680-1750) (Universitat Pompeu Fabra)' and 'Arxivi fotogràfic-Estudis de la Masia Catalana (Centre Excursionista de Catalunya)'. Each result includes a thumbnail image and a brief description.

Presentazione della scheda descrittiva (registrazione attraverso i metadati) dell'oggetto rilevato:

The screenshot shows the descriptive metadata page for the '150th Anniversary / Glasgow Chamber of Commerce and Manufactures'. It includes fields for 'Rating', 'Títol', 'Autor', 'Font', 'Data del document original', 'Paginació', 'Matèria', 'Llengua', 'Editor', 'Tipus de recurs', 'Repositori', and 'Gestió dels drets (Europeana)'. The 'Títol' field is highlighted in blue.

Dichiarazioni sui METADATI

Dichiarazioni su FORMATI

Presentazione del contenuto dell'oggetto con la possibilità di "Text search":

The screenshot shows the document viewer interface. It displays a historical document with a search bar at the top. The document text is visible, and there are navigation controls on the right side.

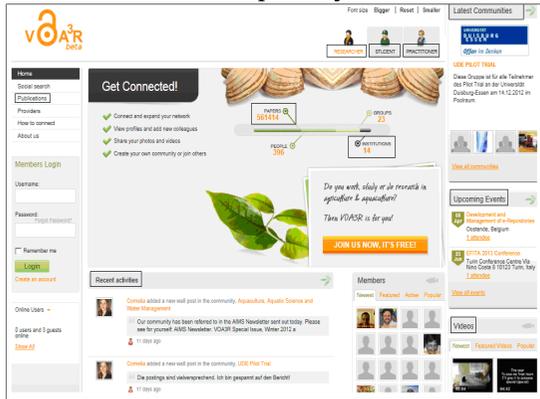
V. Comunicazione mediata dal Repository in collegamento con altri sistemi, moduli:

1. Interazione tra i contenuti delle collezioni di 24 istituzioni cooperanti
2. Condivisione dei contenuti con l'aggregatore delle collezioni storiche sul web "Europeana"
3. A partire dalla scheda descrittiva è possibile:
 - a) condividere le informazioni sull'oggetto tramite vari *Social Networks*;
 - b) valutare (*rating*) e fornire commenti critici sull'oggetto

trovato + non trovato -



I. Presentazione del Repository



Finalità: VOA3R è sviluppato entro un progetto del FP7 - *Research - Europa - ICT Policy Support Program* (2010-2013). Rappresenta un meta-Archivio che - a partire da un unico punto d'accesso - raccoglie, gestisce e diffonde in rete la conoscenza sul materiale di ricerca nel campo di agricoltura prodotto in diverse istituzioni riunite nel **Consorzio VOA3R "European Open Access Association"**

Software pilota disegnato da [RedPanda.gr](http://redpanda.gr) e potenziato da [Joomla!](http://joomla!)¹
Lingua di presentazione: Inglese

Livelli d'accesso:

1. Un **accesso autorizzato** per diversi profili di utenti (ricercatori, studenti, professionisti) affinché possano partecipare in diverse attività comunicative del *Repository*. Un accesso ridirezionato ad altri *Repository* cooperanti per la consultazione diretta di oggetti rilevati tramite VOA3R.

2. Un **accesso libero** per la ricerca e la consultazione di oggetti (*full-text* e/o *Abstract*) depositati e schede di metadati. È possibile scaricare *full-text* se tale non è impedito dal *Copyright* e/o dal periodo di embargo.

Contenuti di ricerca:

1. **Oggetti depositati** (570.000): materiale di ricerca, multimedia
2. **Schede descrittive** (record di metadati) per la rappresentazione di oggetti depositati.

II. Aspetti tecnici per la presentazione di oggetti depositati:

1. **Linking** tra oggetti depositati e profili di autori/utenti
2. **Ultime immissioni:** riepilogo regolarmente aggiornato sui nuovi oggetti depositati tramite le funzioni di *Alerting*, *feed RSS/Atom*.
3. Creazione automatica di file PDF per il **download** di oggetti

III. Navigazione e ricerca dei contenuti:

a) lingua; b) collezioni di *Repository/Provider* partecipativi:

By language	By collection
Arabic (1) Catalan (2)	496737
Czech (22) Danish (1147)	10688
German (3848) Greek (1087)	
English (537434)	

c) **"Timeline view"**; d) **"Drag and drop"**: keywords normalizzate ed espansive dall'AGROVOC Thesaurus e trascinabile sulle sezioni concettuali predefinite.

AGROVOC contiene 40,000 concetti disponibili in 22 lingue ed è pubblicato come *Linked Data* (LD) che interconnette le entità concettuali esponendole sul web e) **keywords** più ricercate: generali e dall'AGROVOC organizzate in *Tags Cloud*; c) **"Map view"**:

IV. Presentazione dei risultati di ricerca e di navigazione:

Presentazione della scheda descrittiva (registrazione attraverso i metadati) dell'oggetto rilevato:

Presentazione del contenuto dell'oggetto rilevato:

V. Comunicazione mediata dal Repository in collegamento con altri sistemi, moduli:

1. Interazione concettuale dei contenuti delle collezioni di 14 *Repository* cooperanti, in base all'impiego dell' *AGROVOC Thesaurus*
2. Funzioni di *Social Networking*: a) *Blog*; b) *Rating*, *annotazioni* e *revisioni* di oggetti depositati, a partire dalle schede descrittive; d) **"Like"** con la condivisione dei contenuti in diversi *Social Networks*; e) **interazione online/offline** tra profili di utenti.
3. Attraverso l'impiego dell' *AGROVOC Thesaurus* pubblicato come *Linked Data*, i contenuti del *VOA3R* possono interagire con altri set di dati affiancati dall' *AGROVOC* nel



POLICY del *Repository*

Dichiarazioni sull'**OPEN ACCESS**

Politiche di editori, **COPYRIGHT**

Dichiarazioni sui **METADATI**

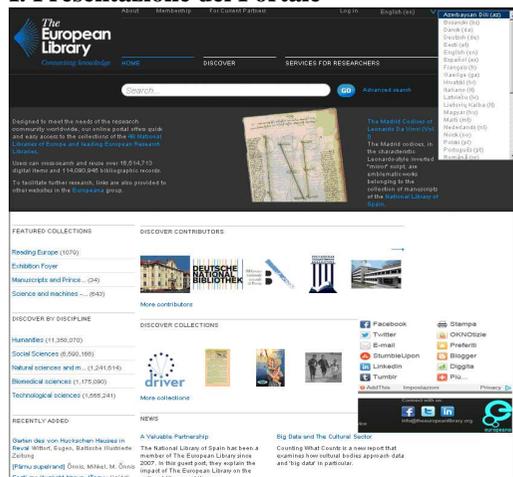
Dichiarazioni su **FORMATI**

trovato + *non trovato* -

¹ <<http://www.redpanda.gr/>; <http://www.joomla.org/>>

Portale „The European Library - Connecting Knowledge” (Olanda) di 48 Biblioteche europee /nazionali e di ricerca:
 <<http://www.theeuropeanlibrary.org/tel4/>>

I. Presentazione del Portale



Software: *Open Source Python*
Lingue di presentazione: 36
Finalità: funzione di un meta-Archivio che raccoglie, collega, diffonde in rete e offre un accesso integrato alla conoscenza sulle risorse bibliografiche digitali/digitalizzati fornite da 48 Biblioteche nazionali europee. Il Portale è il secondo più grande *Content Provider* per l'aggregatore di collezioni culturali "Europeana":

<<http://www.europeana.eu/portal/europeana-providers.html>>

Livelli d'accesso: 1. Accessi autorizzati per:

- a) l'uso delle funzionalità di ricerca avanzate secondo le voci controllate di file di autorità VIAF e GeoNames nella versione *Linked Data*, e MACS;
- b) visualizzare le statistiche di ricerca personali e le voci per l'arricchimento dei metadati dal Portale che accrescono il valore dei metadati raccolti:

<<http://www.theeuropeanlibrary.org/tel4/login?targeturl=%3Flocale%3Dit>>

2. Accesso libero per la consultazione delle descrizioni bibliografiche degli oggetti forniti da diverse istituzioni partecipative.

POLICY del Portale + **Dichiarazioni sull'OPEN ACCESS** + **Politiche di editori, COPYRIGHT** + **Dichiarazioni su METADATI** + **Dichiarazioni su FORMATI** - **trovato** + **non trovato** -

Contenuti di ricerca:

1. **Oggetti digitali/digitalizzati:** 17.000.000
2. **Schede descrittive** (record di metadati, record bibliografici) 115.000.000 record bibliografici per la rappresentazione di oggetti raccolti dall'*Union Catalogue* europeo.

II. Aspetti tecnici per la presentazione di oggetti depositati:

1. **Linking** tra diversi oggetti sulla base di similarità semantica, in conformità con "Europeana Data Exchange Agreement (DEA)":

<<http://pro.europeana.eu/web/guest/data-exchange-agreement>>

2. **Localizzazione geografica** dei contenuti rilevati:



3. Statistiche.

III. Navigazione e ricerca dei contenuti:

1. soluzioni di ricerca integrata "cross-collection" multi-lingua:

Funzionalità avanzate di **Google search**:

2. navigazione per le voci controllate (**Authority search**):

3. navigazione per: contenuti simili di altri contributori; discipline; anni; contributori; collezioni:

IV. Presentazione dei risultati di ricerca e di navigazione:

Presentazione della scheda descrittiva per ogni oggetto rilevato:

Reindirizzamento alla descrizione sul server del contributore:

V. Comunicazione mediata dal Repository in collegamento con altri sistemi, moduli:

L'interazione dei contenuti del *Portale* per similarità con i contenuti di:
 a) "EUROPEANA" (collezioni culturali europee), "CORE" (collezioni da Repositori *Open Access*); *MENDELEY* (una rete collaborativa con le funzionalità di *Social networking* dei contenuti di ricerca):

b) affiancati da VIAF e GeoNames nel

3.4. Risultati dell'indagine

In questo paragrafo sono illustrati i risultati d'indagine svolta per 23 infrastrutture di ricerca (Tabella 10), con l'obiettivo di:

- somministrare le osservazioni particolari e generali;
- evidenziare le migliori funzionalità rilevate entro le piattaforme;
- fornire alcune informazioni sulle funzionalità e strumenti auspicabili per potenziare le funzionalità di un Archivio Istituzionale.

I risultati d'indagine - di cui in Tabella 10 - saranno allestiti secondo le domande distribuite in 5 macroaree in base a cui sono state compilate le schede d'indagine presentate nel paragrafo precedente.

I. Presentazione dell'Archivio (<i>Repository</i>)	
1.	<p>Software</p> <p>1) Quello che accomuna tutte le piattaforme indagate è la loro compatibilità con il protocollo di interoperabilità OAI-PMH (<i>Open Archive Initiative Protocol for Metadata Harvesting</i>) per l'<i>harvesting</i> dei metadati che descrivono le proprietà e il contenuto di oggetti digitali depositati nei <i>Repository</i>. Per mezzo dell'OAI-PMH i medesimi oggetti possono essere ricercati attraverso diversi motori di ricerca</p> <p>a) sia generici come <i>Google, Google Scholar</i>;</p> <p>b) che specialistici del mondo <i>Open Access</i> come <i>Open Access Map, OAIster, WorldCat, Directory of Open Access Repositories (DOAR), Registry of Open Access Repositories (ROAR), Online Computer Library Center (OCLC)</i>, e per i <i>Repository</i> italiani anche PLEADI (Portale per la Letteratura scientifica Elettronica Italiana su Archivi aperti e Depositi Istituzionali) [V.siti].</p> <p>2) I Software sviluppati in conformità con i requisiti dell'<i>Open Access Initiative</i> si presentano sostanzialmente come un mezzo per:</p> <p>a) accelerare la comunicazione tecnica e la condivisione dei contenuti tra diverse piattaforme web;</p> <p>b) aumentarne la visibilità del materiale delle piattaforme e l'efficienza nella disseminazione e ricerca dei contenuti digitali.</p> <p>Sostanzialmente questo è il motivo per cui i Software compatibili con l'OAI-PMH hanno riscosso tanti consensi tra la comunità scientifica e di ricerca, dal loro esordio negli anni Novanta ad oggi.</p> <p>3) È opportuno evidenziare i vantaggi delle piattaforme di Archivi Istituzionali appoggiati sui servizi programmati per la <i>governance</i> dei contenuti di ricerca e dei processi a loro supporto:</p> <p>a) SURplus (Sistema Unificato Ricerca per la gestione delle attività di ricerca Istituzionale) <i>Open Archive</i>, una personalizzazione del Software DSpace;</p> <p>b) <i>UGOV_Ricerca</i>.</p> <p>Gli Archivi Istituzionali AIR (l'Università di Milano) e PORTO (il Politecnico di Torino) - sviluppati per mezzo dei succitati servizi - vengono direttamente collegati con le funzioni dell'Anagrafe della ricerca delle Istituzioni per cui sono implementati e, di conseguenza, anche con le informazioni che riflettono gli adempimenti connessi al sistema di valutazione della ricerca Ministeriale.</p> <p>La documentazione che constata la tendenza della connessione degli Archivi Istituzionali del materiale di ricerca con gli Anagrafi della ricerca:</p> <p>a) "<i>Rome Declaration on CRIS and OAR</i>" [V. sito];</p>

	<p>b) <i>“State of the Art and Roadmap for Current Research Information Systems and Repositories”</i> [V. sito];</p> <p>Nel 2013 in Inghilterra viene ampiamente ribadita la centralità degli Archivi Istituzionali <i>Open Access</i>, non solo per la disseminazione centralizzata dei contenuti di ricerca, ma anche per l'esercizio nazionale di valutazione della ricerca. Al fine della medesima valutazione verrebbero accettati solo i contributi depositati negli Archivi Istituzionali <i>Open Access</i>, come segue dalla proposta HEFCE (<i>Higher Education Funding Council England</i>) per il prossimo esercizio nazionale di valutazione della ricerca <i>Research Excellence Framework (REF)</i>:</p> <p><i>To support and encourage the further implementation of Open Access we intend to introduce a requirement that all outputs submitted to the post-2014 Research Excellence Framework (REF) exercise are published on an Open-Access basis. As part of our commitment to increasing public access, we intend to require that outputs meeting the REF Open Access requirement (whether published by the “Gold” or “Green” route) shall be accessible through an Institutional Repository. This reflects our view of the significant role of Institutional Repositories in increasing sustainable and convenient public access to research [V. sito “Research Excellence Framework (REF)].</i></p>
	<p>Finalità</p>
	<p>1) La maggior parte di Archivi digitali, su cui è stata eseguita l'indagine, ha il fine di raccogliere e disseminare tramite Internet il materiale di ricerca prodotto entro le rispettive Istituzioni.</p> <p>2) 8 tra 22 <i>Repository</i>:</p> <ul style="list-style-type: none"> ✓ <i>Phaidra</i>, ✓ <i>Cambridge University DSpace</i>, ✓ <i>CERN Document Server</i>, ✓ <i>Archive Ouverte HAL-INRIA</i>, ✓ <i>Minho's Institutional Repository</i>, ✓ <i>MIT DSpace</i>, ✓ <i>Dipòsit Digital de Documents de la UAB</i>, ✓ <i>UPCommons</i> <p>ospitano al loro interno anche il materiale didattico e alcuni tra loro anche il materiale amministrativo.</p> <p>3) 5 tra 22 <i>Repository</i>:</p> <ul style="list-style-type: none"> ✓ <i>Cambridge University DSpace</i>, ✓ <i>Minho's Institutional Repository</i>, ✓ <i>MIT DSpace</i>, ✓ <i>Open Repository and Bibliography ORBi</i>, ✓ <i>Europe PubMed Central Subject Repository</i> <p>raccogliono al loro interno anche i dati di ricerca (<i>“Research Data”</i>¹);</p> <p>4) 2 <i>Repository</i> italiani AIR e PORTO sono collegati ai moduli per la valutazione dei risultati di ricerca accademica;</p> <p>5) 6 tra 22 <i>Repository</i>:</p>

¹ *“RESEARCH DATA means data in the form of facts, observations, images, computer program results, recordings, measurements or experiences on which an argument, theory, test or hypothesis, or another research output is based. Data may be numerical, descriptive, visual or tactile. It may be raw, cleaned or processed, and may be held in any format or media [...] Research Data, from the point of view of the Institution with a responsibility for managing the data includes: (1) all data which is created by researchers in the course of their work, and for which the institution has a curatorial responsibility for at least as long as the Code and relevant archives/record keeping acts require, and (2) third-party data which may have originated within the institution or come from elsewhere”,*
Australian National Data Service: <<http://ands.org.au/guides/what-is-research-data.html>>

	<ul style="list-style-type: none"> ✓ Phaidra, ✓ Archive Ouverte HAL-INRIA, ✓ MDC: Digital Memory of Catalonia, ✓ OpenAIRE Orphan Record Repository, ✓ Europe PubMed Central Subject Repository, ✓ UPCommons, <p>insieme al Portale di ricerca “The European Library” svolgono la funzione di c.d. “Meta-Archivi” in quanto raccolgono e diffondono in rete - a partire da un unico punto d’accesso - le informazioni sul materiale depositato in diversi Archivi e Biblioteche digitali cooperanti, secondo gli accordi sviluppati.</p>
	<p>Livelli di accesso</p>
	<p>1) La maggior parte delle piattaforme oggetto d’indagine prevede le seguenti modalità di <i>accesso</i>:</p> <p>a) <i>autorizzato</i> per:</p> <ul style="list-style-type: none"> ✓ il deposito (<i>upload</i>) del materiale da parte del personale coinvolto nella ricerca, degli studenti e/o del personale amministrativo delle rispettive Istituzioni; ✓ la consultazione di contenuti di determinate categorie di oggetti depositati; ✓ la visualizzazione delle statistiche; ✓ la valutazione di oggetti depositati tramite i moduli di “<i>Rating</i>”, “<i>Comments</i>”, “<i>Discussions</i>”, a partire dalle scheda descrittive di oggetti depositati; <p>b) <i>libero</i> per la ricerca e consultazione di:</p> <ul style="list-style-type: none"> ✓ <i>full-text</i> e/o ✓ <i>Abstract</i>; ✓ schede di metadati (sommario e metadati estesi) <p>relativi agli oggetti depositati.</p> <p>Inoltre, in ogni <i>Repository</i> è possibile scaricare <i>full-text</i> se ciò non è impedito dal <i>Copyright</i> e/o dal periodo di embargo (di solito tra 6-12 mesi) o se non ci sono <i>accessi ristretti</i> (es. “<i>Campus access</i>”, “<i>For archival purposes only</i>”) che possono riguardare i determinati gruppi di utenti.</p> <p>2) La presenza di accessi diversificati sono rilevati nel <i>Repository</i> PHAIDRA:</p> <p>a) <i>accessi autorizzati</i> per il <i>caricamento</i> e <i>gestione</i> dei <i>contenuti</i> assegnati ai tre gruppi di utenti: “<i>Administrator</i>”, “<i>Super-User</i>”, “<i>Owner</i>”;</p> <p>b) <i>accessi per la ricerca dei contenuti</i> assegnati ai quattro gruppi di utenti: “<i>Administrator</i>”, “<i>Super-User</i>”, “<i>Owner</i>”, “<i>Anonyme User</i>”.</p> <p>3) Nel Portale “<i>The European Libraries</i>” gli <i>accessi autorizzati</i> sono assegnati agli utenti che intendono:</p> <p>a) fare l’uso delle funzionalità di ricerca avanzate tra cui la ricerca dei contenuti secondo le voci dei vocabolari controllati [Harpring, Baca, 2010] come <i>Virtual International Authority File</i> (VIAF) [V. sito] che controllano i nomi di persone e di istituzioni identificati dagli indirizzi univoci web URI; <i>GeoNames</i> [V. sito] che rende disponibili più di 6 milioni di toponimi; MACS (<i>Multilingual Access to Subject</i>) che è un authority file virtuale multilingue [V. sito];</p> <p>b) visualizzare le statistiche di ricerca personali e le voci per la qualificazione delle proprietà di metadati che accrescono il valore di questi ultimi [V. sito “<i>The European Libraries - qualificazione metadati</i>”].</p>
	<p>Policy del <i>Repository</i> (trovato +/- non trovato -)</p>

	<p>1) Quasi tutti i <i>Repository</i> sono allestiti con le Politiche/<i>Policy</i> in cui vengono dichiarati le loro funzioni, l'uso, le strutture organizzative e le finalità.</p> <p><i>Every Repository should ideally have Policies on:</i></p> <ul style="list-style-type: none"> - <i>the re-use of metadata</i> - <i>the re-use of full texts (or their equivalent)</i> - <i>permitted Content - i.e. the subjects and types of item that can be deposited</i> - <i>Submission Requirements, including moderation of new items</i> - <i>Long Term Preservation of the data.</i> <p><i>Policies should be visible to users as downloadable documents or web pages in your Repository</i> [V. sito: "OpenDOAR_Help"].</p> <p>2) Tutti i <i>Repository</i> sono allestiti con le Politiche relative all'immissione/deposito di oggetti digitali (con la specifica di loro tipologie). Intanto:</p> <p>Per facilitare l'immissione dei contenuti in un Archivio Istituzionale viene raccomandata "una maschera di immissione semplice ed intuitiva, magari con una procedura guidata passo per passo, a portata di click [Guerrini, 2010, p.18].</p> <p>Gli aspetti che possono essere migliorati sono:</p> <p>1) la maggioranza dei <i>Repository</i> dovrebbero provvedere ad un miglior allestimento delle Politiche su:</p> <ul style="list-style-type: none"> ✓ metadati implementati; ✓ riuso di metadati e di <i>full-text</i>; ✓ conservazione a lungo termine. <p>2) Nell'allestimento delle Politiche è di cruciale importanza la valorizzazione del deposito dei lavori <i>Open Access</i> e di introdurre l'uso delle metriche di valutazione rese possibili dagli strumenti basati su <i>Social Media</i>:</p> <p><i>Policy-makers and University administrators begin to formulate Policy and framework that would "move prestige to Open Access" by rewarding publications in high-quality Open Access outlets and by recognizing the broader range of impact metrics made possible by social media</i> [V. sito: "In defence of Open Access systems"].</p> <p>3) È opportuno che le Istituzioni optino per lo sviluppo di Politiche istituzionali a supporto del deposito obbligatorio ("<i>Mandatory Policies</i>") del materiale di ricerca svolte nell'ente nei rispettivi Archivi Istituzionali,</p> <p>come <i>condizione per accedere ai relativi finanziamenti</i>. A ben vedere, questo tipo di regolamenti non fissa un vero e proprio obbligo per l'autore, piuttosto definisce un onere cui questi deve conformarsi se vuole ottenere un beneficio (accesso a un finanziamento; ammissione alla discussione di una tesi per il completamento di un ciclo di studi svolto avvalendosi di strutture e risorse dell'ente).</p> <p>L'autore rimane arbitro assoluto della decisione sul deposito, tanto che potrà provvedere egli stesso all'auto-archiviazione corredandola di apposita Licenza d'uso destinata al pubblico (CC o altro), ovvero l'Istituzione gli chiederà di sottoscrivere una Licenza (il cui destinatario è la stessa istituzione e ove vengono regolati i rapporti con l'autore) e solo a seguito di quest'autorizzazione provvederà al deposito [...] È auspicabile quindi il diffondersi di Politiche Istituzionali di questo tipo e, meglio ancora, l'affermazione del principio dell'Accesso Aperto a livello normativo [Guerrini, 2010, pp.123-124].</p>
	Dichiarazioni sull'OPEN ACCESS (trovato +/- non trovato -)
	Quasi tutti i <i>Repository</i> sono allestiti con le informazioni che descrivono gli aspetti

	<p>principali del movimento <i>Open Access</i> a supporto della disseminazione dei risultati della ricerca in rete.</p> <p>Questo requisito è di primaria importanza in quanto la maggioranza dei <i>Repository</i> sono originati proprio in vista della promozione del materiale di ricerca in modalità <i>Open Access</i>.</p>
	<p>Politiche di Editori, COPYRIGHT (trovato +/- non trovato -)</p>
	<p>1) Quasi tutti i <i>Repository</i> sono allestiti con le informazioni relative alle Politiche di Editori e alle questioni di <i>Copyright</i>, per informare gli utenti sulle condizioni del deposito e/o delle modalità di consultazione dei contenuti digitali.</p> <p>L'attenzione nei confronti delle questioni di <i>Copyright</i> - per le quali l'Ateneo di Padova (Italia) ricopre un ruolo guida - dovrebbero contribuire a fugare i timori degli autori verso la pratica dell'auto-archiviazione negli Archivi Istituzionali/Disciplinari di ricerca [V. sito: "Diritto d'autore_Università di Padova".</p> <p>2)</p> <p>L'aspetto dei Diritti comporta un'attenta riflessione, in quanto il processo di deposito, disseminazione e fruizione di un lavoro entro un Archivio Aperto è permeato in tutte le fasi dalla delicata gestione dei Diritti. La <i>gestione è assai diversa</i> se si tratta di <i>materiale di ricerca</i> o di <i>materiale didattico</i>. Mescolare queste due tipologie di documento in un unico IR potrebbe risultare <i>problematico</i>, sia in termini di <i>comunicazione</i> su come e cosa trattare in merito ad <i>autorizzazioni e liberatorie</i>, sia in termini di chiarezza nella <i>definizione dell'identità di un IR</i> [Guerrini, 2010, p.30].</p> <p>3) Secondo il documento "<i>In defence of Open Access systems</i>" [V. sito], molti <i>Repository</i> Istituzionali rimangono scarsamente popolati in quanto spesso le loro pagine di presentazione dedicano poca attenzione per mettere in risalto le informazioni sui vantaggi dell'<i>Open Access</i>, sulla qualità del materiale <i>Open Access</i> compatibile con il processo di "<i>peer review</i>", sulle questioni di <i>Copyright</i>.</p>
	<p>Contenuti di ricerca rilevati</p>
	<p>1) Nella maggior parte dei <i>Repository</i> Istituzionali i contenuti vengono rappresentati da:</p> <ul style="list-style-type: none"> a) articoli, <i>e-books</i> (libri digitali/digitalizzati) e/o loro capitoli, monografie, rapporti tecnici, interventi in convegni e conferenze, <i>working papers</i>, Tesi di Dottorato. Tra tutte queste tipologie c'è la preponderanza di articoli di ricerca - che è il mezzo di comunicazione per eccellenza della comunicazione scientifica - e di tesi di dottorato; b) <i>schede bibliografiche</i> (record di metadati) che descrivono gli oggetti di cui nel punto (a). <p>4) Alcuni <i>Repository</i> predispongono i loro spazi anche per il deposito del materiale:</p> <ul style="list-style-type: none"> a) multimedia (film, filmati audio e video, audio); b) collezioni di fotografie e immagini; c) il materiale didattico (incluso quello ad accesso libero "<i>Open Courseware</i>"); d) il materiale amministrativo; e) dati di ricerca che possono includere anche i così detti "dati grezzi" (<i>raw data</i>) i quali devono essere depositati in formati standard per facilitare il loro riconoscimento da parte di sistemi informatici e le eventuali operazioni di <i>text-</i> e <i>data-mining</i>, le simulazioni gli studi trasversali sui dati.

Gli Atenei sono chiamati, in modo sempre più pressante, ad erogare servizi online per la propria utenza. L'offerta spazia dai Portali didattici ai *Repository* della ricerca, sino alle informazioni relative all'offerta didattica rivolte agli studenti [V. sito: "CINECA"].

5) L'apporto al materiale didattico è dato maggiormente dai *Repository* dedicati, creati per dare più rilevante visibilità al medesimo materiale valorizzato come un prodotto della ricerca scientifica di qualità, spesso poco promosso, e per servire da supporto all'attività didattica.

6) Nel caso un'Istituzione decidesse di raccogliere e valorizzare i materiali digitali di diverse tipologie, sarebbe opportuno creare diversi sotto-Archivi separati (es. un sotto-Archivio dedicato alle Tesi di Dottorato, un altro – al materiale didattico, altri ancora - agli articoli e pubblicazioni, alle immagini, ai video e così via).

L'importanza di assegnare diverse tipologie di oggetti informativi digitali a diverse categorie o sotto-Archivi potrebbe essere giustificata dal fatto che le varie tipologie del materiale dovranno, di buona norma, essere affiancati da diversi metadati e Diritti per la propria gestione, da diverse Politiche per la consultazione e la conservazione. Queste prospettive gestionali potrebbero essere indubbiamente impegnative per un'Istituzione:

Se infatti già la distinzione fra strumenti LCMS [*Learning Content Management System*] e OA [*Open Access*] pone notevoli problemi di sostenibilità gestionale, cosa dire di una situazione in cui le tipologie di *Repository* richieste a una Istituzione che desidera archiviare e distribuire i propri materiali didattici e di ricerca possono essere addirittura tre (OA, LCMS e piattaforma specifica per l'OCW [*Open CourseWare*])? Inoltre, l'uso di una piattaforma totalmente indipendente per la gestione del *CourseWare* non risponde al requisito di una buona integrazione con la gestione dei prodotti della ricerca (questa integrazione potrebbe essere in parte offerta a livello di Service Provider sfruttando il supporto per il protocollo OAI-PMH, il che però comporterebbe il ricorso a un'ulteriore tipologia di strumenti [Guerrini, 2010, p.29].

Alcune Università sviluppano infatti Archivi dedicati esclusivamente ai *prodotti della ricerca*, altre estendono invece il servizio anche ad attività istituzionali come la *didattica*; altre ancora creano più di un Archivio di deposito, per tenere *distinte*, per esempio, la *ricerca* dalla *didattica* e dalle *Tesi di Laurea* e di *Dottorato* [ivi., p.144].

La ricerca è (dovrebbe essere) unita alla didattica e risultati di una ricerca in corso (e tanto più se conclusa) sfociano spesso in una dispensa o in un testo nel quale i concetti della ricerca vengono proposti in un linguaggio didattico. Per questa ragione è importante che i materiali didattici guardino verso l'*Open Access* e siano realizzati secondo questa concezione [secondo due strategie]. La prima è la creazione di Archivi Aperti che contengano materiale didattico di varia tipologia, in modo distinto rispetto all'Archivio in cui sono depositati i prodotti della ricerca. In questa organizzazione si configurerebbero tre ambienti:

- (1) l'IR preposto alla ricerca,
- (2) un Archivio per il materiale didattico e
- (3) la piattaforma a distanza, contenente i *Learning Object* (LO) che non possono, per loro natura, risiedere fuori dalla piattaforma nativa [ivi., p.29].

Una simile configurazione dell'albero delle collezioni rappresenta l'Archivio Istituzionale dell'Università di Bergamo AISBERG, cui piattaforma ospita:

1. una sezione per le pubblicazioni per la valutazione della ricerca;
2. una sezione per "*Scholarly publishing initiatives*";
3. una sezione per "*Open educational resources*".

	[V. sito: “AISBERG, ovvero, come far convivere pacificamente la valutazione della Ricerca con la pubblicazione OA in un unico Archivio”].
II. Aspetti tecnici a supporto di presentazione di oggetti depositati	
	Dichiarazioni sui METADATI a supporto dei contenuti (trovato +/- non trovato -)
	<p>1) La maggioranza dei <i>Repository</i> è allestita con le informazioni sui metadati che descrivono le proprietà e il contenuto degli oggetti depositati.</p> <p>2) Diversi <i>Repository</i> forniscono tramite le loro pagine introduttive i rimandi per la consultazione delle informazioni sui Metadati collocate sui siti di Software per mezzo di cui i medesimi <i>Repository</i> sono implementati. È opportuno che tutti i <i>Repository</i> esponano a sufficienza le informazioni sugli schemi di metadati e sulle mappature che essi adottano, considerando anche le eventuali personalizzazioni a seconda delle particolarità delle proprie tipologie di oggetti da depositare.</p>
	Dichiarazioni sui formati (trovato +/- non trovato -)
	La maggioranza dei <i>Repository</i> viene allestita con le Politiche che dichiarano le informazioni sui formati a supporto degli oggetti depositati.
	Assegnazione di indirizzi/identificatori persistenti agli oggetti depositati
	<p>1) La maggioranza dei <i>Repository</i> assegna ad ogni oggetto depositato indirizzi unici e persistenti, che garantiscono l’accesso ai contenuti a lungo termine entro il <i>Repository</i> di riferimento. Tra tali indirizzi prevalgono gli identificatori <i>HANDLE system® persistent identifier</i> e <i>DOI</i>. Tra altri possibili indirizzi persistenti vi sono anche <i>URN</i>, <i>BiTex entry</i>, <i>PURL</i>, <i>MD5 PGP keys</i>, <i>ISBN</i>.</p> <p>2) L’assegnazione di un identificatore univoco e persistente come, ad esempio, previsto nativamente dalle piattaforme (Software) <i>DSpace</i> e <i>Fedora</i>, la preferenza accordata all’archiviazione in <i>formati standard</i> che garantiscono la preservazione degli oggetti digitali, sono tutti indici di attenzione verso il tema della conservazione dei contenuti a lungo termine.</p>
	Memorizzazione di <i>stream</i> di video e audio
	I <i>Repository</i> che prevedono il deposito anche dei file multimedia hanno i meccanismi per la memorizzazione di <i>stream</i> di video e audio.
	Statistiche d’uso (accesso, visualizzazione, <i>download</i>)
	<p>1) La maggior parte di <i>Repository</i> sono allestiti con le funzionalità che permettono di visualizzare le statistiche d’uso riguardanti il numero di accessi e di <i>download</i> di oggetti e collezioni di oggetti [V. “<i>Using usage statistics to encourage deposits</i>”].</p> <p>2) Una buona pratica per diversificare la rappresentazione delle statistiche d’uso è fornita dal <i>Repository</i> “<i>MIT DSpace</i>”. Esso fornisce le statistiche di <i>download</i> del materiale di ricerca <i>Open Access</i> prodotto entro il MIT, distribuito per i continenti:</p>

	
	<p>Funzionalità di <i>upload</i> (caricamento), l'importazione e l'esportazione dei contenuti</p>
	<p>Un <i>Repository</i> di qualità deve fornire la possibilità all'utente di caricare e/o importare/esportare, esportare non solo i singoli file, funzionalità prevista da tutti gli Archivi, ma anche un insieme di file, mega-file (intere collezioni) - con un solo <i>upload</i>, come viene offerto dal <i>Repository</i> PHAIDRA tramite le funzionalità "<i>Bulk-Uploads</i>", "<i>Book Importer</i>", "<i>Phaidra importer</i>".</p>
	<p>Funzionalità di <i>alerting</i> e <i>feed RSS/Atom</i> per l'aggiornamento sui nuovi oggetti depositati</p>
	<p>La maggior parte di <i>Repository</i> sono allestiti con funzionalità che permettono di:</p> <ol style="list-style-type: none"> ricevere aggiornamenti via <i>e-mail</i> (<i>alerting</i>) sugli oggetti recentemente depositati; visualizzare sulle loro pagine le informazioni su ultimi contenuti depositati ("<i>Most recent</i>", "<i>Recently added</i>") attraverso le funzioni di <i>feed RSS/Atom</i>.
<p>III. Ricerca, navigazione/<i>Browsing</i> dei contenuti</p>	
	<p>Funzionalità di ricerca e di navigazione/<i>Browsing</i></p>
	<p>1) Tutti i <i>Repository</i> vengono allestiti con le funzionalità "Ricerca Semplice" e "Ricerca Avanzata".</p> <p>Il blocco "Ricerca Avanzata" ("<i>Advanced Search</i>") è composto da una serie di box su cui sono esplicitati tutti i filtri disponibili per la ricerca avanzata. Occupa l'intera pagina cui e' posizionato. Sommarariamente, i filtri implementati sono:</p> <ul style="list-style-type: none"> • tre livelli di criteri di ricerca sovrapponibili (Parola Chiave/ Nome/ Soggetto/ <i>full-text</i>) con operatori diversi (Inizia con/ Frase esatta/ Tutte le parole/ Qualsiasi parola) e operatori Booleani (AND, OR, NOT); • tipo di risorsa digitale; • collezione di appartenenza; • data; • <i>accesso pubblico</i> (basato su una serie di metodi e funzionalità che sono ritenuti per la loro natura "<i>pubblici</i>" e quindi non richiedono l'autenticazione dell'utente; • <i>accesso ristretto</i>. <p>Di buona norma i risultati delle ricerche:</p> <ul style="list-style-type: none"> • devono essere filtrati (tramite l'applicazione di un filtro/una sorta di <i>proxy</i> lato <i>Drupal</i> sulle <i>query</i>) in base alle richieste provenienti da un determinato range di IP, ai ruoli e ai diritti di accesso che ha l'utente correntemente loggato (o l'utente anonimo per i contenuti pubblici); • vengono prodotti dalla stessa pagina della ricerca semplice e quindi possono

essere ulteriormente raffinati utilizzando le faccette impostate nella configurazione e la ricerca *full-text* nei risultati. Cliccando sui link che affiancano le faccette restringe i risultati per quella faccetta.

2) Tutti i *Repository* sono allestiti anche con le funzionalità di navigazione/*Browsing* tramite i *link* a supporto delle parole chiave predefinite (libere e/o controllate).

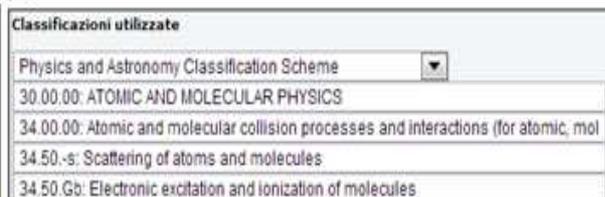
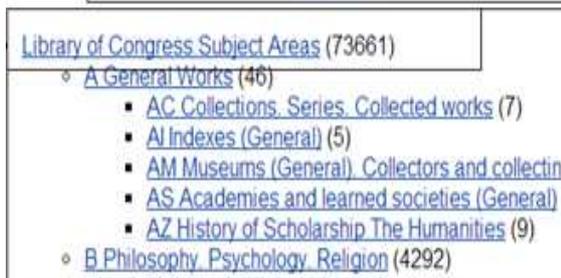
3) La piattaforma del Repository PHAIDRA implementata sul *Software Fedora* contiene il blocco di navigazione che elenca i tipi di oggetti digitali disponibili. Cliccando su uno di essi si scatena la funzionalità di ricerca che mostra tutte le risorse digitali appartenenti a quella tipologia.

4) Le piattaforme dei Repository implementate sul *Software Dspace* e la piattaforma “*The European Library*” dispongono dei blocchi “*Collection Browser*” nella regione che ospita il contenuto principale. In questi blocchi vengono mostrate le schede brevi di tutte le collezioni delle piattaforme. Sotto il link di ogni collezione l’utente può filtrare le sotto-collezioni in tempo reale tramite le classi di faccette predeterminate. Un clic su queste ultime filtrerà le collezioni nella pagina.

Semantica e autorità delle voci di navigazione

1) In diversi Archivi l’attribuzione del *valore aggiunto* per la *navigazione* delle strutture concettuali gerarchiche a *faccette* (o *multidimensionale*) e/o *enumerative*: avviene attraverso l’implementazione (il caricamento sui Server) di:

- Thesauri*, anche di quelli sviluppati localmente;
- Classificazioni (Disciplinari/Tematici e/o di Soggetto) e
- altri vocabolari controllati impiegandoli per la navigazione.



2) L’implementazione nei *Repository* di specifici sistemi di organizzazione della conoscenza (*Knowledge Organisation Systems/KOS*) come Classificazioni, *Thesauri* offre la possibilità di:

- creare la funzione “*navigazione semantica a faccette*”,
- affiancare la ricerca dei contenuti con gli indici (*keywords*) nominali identificativi aventi sufficienti forme di controllo sulle proprietà di metadati che denotano “Soggetto”, “Autore”, “Ente”, “Area geografica” ecc.;

c) navigazione più potenti in quanto più raffinati e autorevoli.

3) Con l'impiego di determinati *KOS* entro le piattaforme digitali di ricerca:

- a) vengono supportate le attività contro la proliferazione di errori e di ambiguità semantica a cascata nella indicizzazione (metadattazione) dei contenuti;
- b) viene risolto facilmente il problema degli omonimi e dei sinonimi con la ricaduta apprezzabile sull'attendibilità dei dati citazionali.

4) Una buona pratica dell'impiego del *Thesaurus* per la ricerca e navigazione dei contenuti fornisce il *Repository VOA3R*, su cui Server viene implementato il *Thesaurus AGROVOC* (40,000 concetti disponibili in 22 lingue) pubblicato *online* tramite la tecnologia *Linked Data*. Le voci controllate semantiche del medesimo *Thesaurus* – identificate univocamente tramite gli indirizzi web *http://URIs* - vengono impiegate per:

- a) la qualifica dei valori di metadati (indici/*keywords*) a supporto descrittivo delle proprietà e del contenuto di oggetti depositati.
- b) la funzione di navigazione "*DDRAG and DROP*" (Figura 18), che permette all'utente di applicare (trascinare) i *keywords* inseriti per la ricerca sui campi semantici predefiniti, costituiti sulle voci controllate dell'AGROVOC.



Figura 18. La funzione di navigazione "*DDRAG and DROP*" basata sui campi concettuali definiti dal Thesaurus AGROVOC

c) Popolare le sezione di ricerca "*Map View*" e "*Tags Cloud*" (Figura 19).



Figura 19. Le sezioni di ricerca "*Map View*" e "*Tags Cloud*" potenziate dalle voci del Thesaurus AGROVOC

Il blocco "*Tags Cloud*" è realizzato tramite le faccette dei temi cliccabili formattata secondo lo standard del *Tagcloud*: la grandezza del font di un termine è proporzionale alla frequenza relativa di quel termine nel *tagging* delle collezioni e delle risorse. I *tag* sono collegamenti che portano alla pagina dei risultati della ricerca dove è stato applicato il filtro in base al tag cliccato.

4) Considerando che il *Thesaurus AGROVOC* è pubblicato come *Linked Data*, esso potrebbe:

a) rappresentare uno strumento mediatore tra i contenuti di un *Repository* e i contenuti di altre piattaforme di informazioni digitali affiancati dalle voci di AGROVOC;

b) integrare i contenuti di *Repository* nel *Linked Open Data Cloud* (Figura 20) che presenta il panorama di interconnessione di una diversità di *dataset* pubblicati sul web come *Linked (Open) Data* da diversi soggetti produttori dei contenuti digitali.

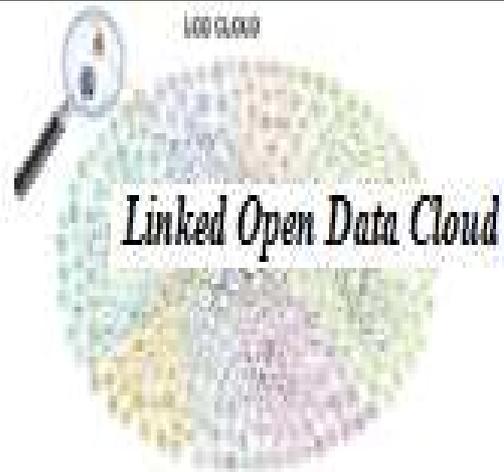


Figura 20. *Linked Open Data Cloud* [V.sito]

5) Nella sezione di ricerca e navigazione del Portale “*The European Library*” - “*Authority search*” (Figura 21) - vengono impiegati i seguenti vocabolari controllati pubblicati online come *Linked Data*:

a) *Virtual International Authority File* (VIAF) [V.sito] per il controllo dei nomi (di persone e di enti) identificati dagli indirizzi univoci web;

b) *GeoNames* [V.sito] che identifica tramite <http://URIs> più di 6 milioni di toponimi;

c) *MACS (Multilingual Access to Subject)* [V.sito].

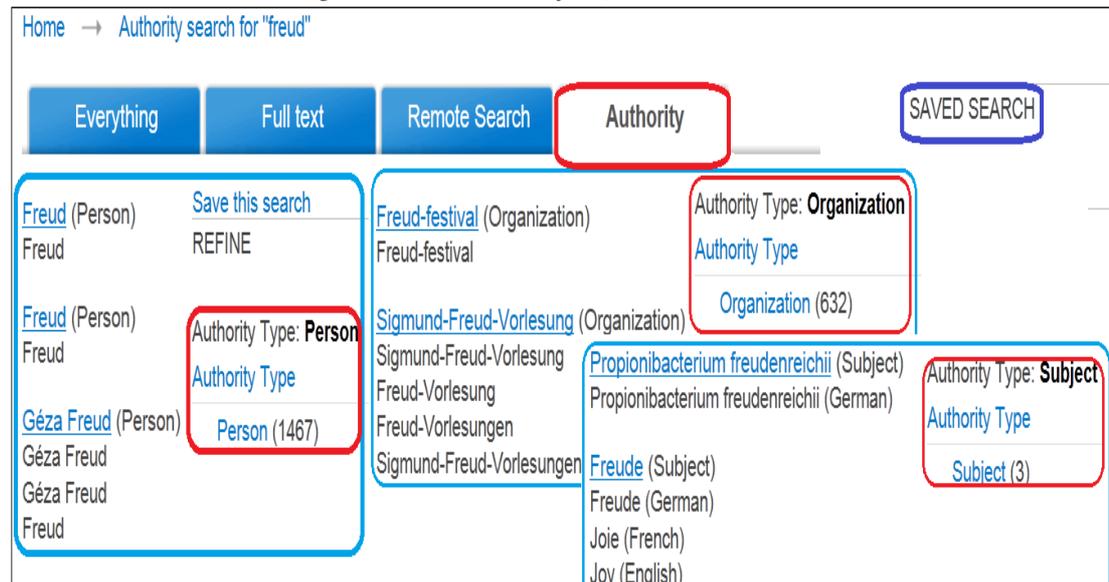


Figura 21. La sezione di ricerca e navigazione del Portale “*The European Library*” - “*Authority search*”

6) Con l’implementazione di KOS è possibile sviluppare strutture di ricerca e navigazione interattive tramite cui “*the visualization of Knowledge can help advance the application of documentary and bibliographic classifications in information and Knowledge discovery*” [V.sito: “*Classification & Visualization: Interfaces to Knowledge*”] come, ad esempio:

a) “*Dynamic Knowledge*”,

	<p>b) “<i>Search Expansion</i>”, c) “<i>Interactive searching</i>”.</p> <p>7) Nel 2013 è uscita la versione dell’UNESCO Thesaurus [V.sito] pubblicata in modalità <i>Linked Data</i>, la cui nomenclatura può essere a supporto della descrizione semantica dei contenuti nel campo umanistico.</p> <p>8) Si fa riferimento ai vocabolari controllati e quelli semantici, pubblicati in modalità <i>Linked (Open) Data</i> e registrati nella banca dati online <i>CKAN Registry (Comprehensive Knowledge Archive Network)</i> [V.sito], e ai casi d’uso nel <i>Library Linked Data Report (2011)</i> [V.sito].</p> <p>9) Dal momento che sovente gli articoli accademici vengono pubblicati da più autori, si rende necessario, per una loro corretta identificazione, predisporre un identificatore univoco che possa collegare le pubblicazioni agli autori. A soddisfare tale requisito e per risolvere il problema di ambiguità dei nomi, il JISC (<i>Joint Information Systems Committee</i>) ha accolto favorevolmente e sta promuovendo l’adozione il sistema per l’identificazione univoca degli autori ORCID: “<i>Open Researcher and Contributor Identifier</i>” [V.sito: “<i>Open-Researcher-Contributor ORCID</i>”]. Se utilizzato su larga scala, ORCID consentirà di:</p> <ul style="list-style-type: none"> a) attribuire a ciascun autore le proprie pubblicazioni, oltre a rendere possibile una ricerca più accurata degli articoli; b) collegare informazioni tra Istituzioni e sistemi a livello internazionale (per quanto riguarda Editori, Finanziatori della ricerca, Istituzioni).
	<p>IV. Presentazione (visualizzazione) dei risultati di ricerca e di navigazione In questa sezione saranno raggruppate le informazioni che trattano gli aspetti:</p> <ul style="list-style-type: none"> a) riguardanti le modalità di presentazione dei contenuti nelle piattaforme su cui è stata eseguita l’indagine; b) auspicabili per migliorare le modalità di visualizzazione dei contenuti.
	<p>Elenco di contenuti recuperati tramite le funzionalità di ricerca e navigazione/<i>Browsing</i></p>
	<p>Dopo la procedura di ricerca e navigazione, ogni piattaforma genera gli elenchi di oggetti recuperati, in conformità con le modalità di <i>layout</i> impostate entro i Software su cui sono implementate le rispettive piattaforme.</p>
	<p>Schede descrittive (record di metadati) per gli oggetti rilevati</p>
	<p>1) Dopo la rilevazione dell’oggetto ricercato, ogni piattaforma prevede la possibilità di consultazione della sua scheda descrittiva, costituita da indici (metadati o dati bibliografici) che descrivono le proprietà e i tratti riassuntivi del contenuto dell’oggetto in questione.</p> <p>2) Si nota la presenza di piattaforme che, nonostante siano implementate sugli stessi Software, presentano i <i>record</i> di metadati (schede descrittive) con diverso numero di elementi. Ciò dipende dalle scelte fatte per la descrizione dei contenuti. Alcune schede riportano gli elementi dei metadati (es. “<i>Author</i>”, “<i>Subject</i>”, “<i>Similar Items</i>”) affiancati da <i>link</i> di approfondimento, cliccando sui quali è possibile risalire alle ulteriori pagine di dettaglio. Ad esempio, il blocco “<i>Similar Items</i>” racchiude l’elenco delle risorse simili alla risorsa correntemente visualizzata nel corpo della pagina.</p> <p>3) La maggior parte delle piattaforme presenta l’allestimento delle schede descrittive di</p>

	<p>oggetti recuperati con:</p> <ul style="list-style-type: none"> a) link di sharing sui <i>Social Network</i>; b) gli spazi per i commenti (<i>Discussions, Comments</i>) critici e per il <i>Rating</i> dei contenuti recuperati; c) i contatti diretti di autori degli oggetti e i <i>link</i> che reindirizzano l'utente per la consultazione di Curriculum (CV) degli autori. <p>5) I <i>Repository</i>:</p> <ul style="list-style-type: none"> ✓ Cambridge University DSpace, ✓ Archive Ouverte HAL-INRIA, ✓ Minho's institutional Repository, ✓ Munich Personal RePEc Archive <p>offrono le funzionalità dell'importazione automatica delle citazioni (liste di pubblicazioni) nei CV personali degli autori depositanti [V. sito: "<i>Automated downloading of citation data</i>".]</p>
Visualizzazione del contenuto dell'oggetto rilevato	
	<p>1) La maggioranza delle piattaforme su cui si è svolta l'indagine fornisce i <i>layout</i> per la consultazione degli oggetti recuperati tramite ricerca e/o browsing, a seconda delle impostazioni <i>default</i> dei rispettivi Software;</p> <p>2) Le schede descrittive del maggior numero di <i>Repository</i> presentano gli elementi di <i>Thumbnail</i> (un'anteprima di un'immagine più grande, che quindi viene presentata in formato ridotto) e i relativi collegamenti, cliccando sui quali vengono aperti diversi tipi di visualizzatore nel browser corrispondenti ai diversi tipi di risorsa e/o download se previsto. Il servizio di visualizzazione e le sue caratteristiche sono codificate nel <i>Repository</i> all'interno del <i>Content Model</i> con cui l'oggetto digitale viene rappresentato;</p> <p>3) I <i>Repository</i>:</p> <ul style="list-style-type: none"> ✓ PHAIDRA; ✓ Open Access Institutional Repository Cadmus; ✓ MDC: Digital Memory of Catalonia <p>sono allestiti con le funzionalità di visualizzazione dei contenuti come: "<i>Book Viewer</i>" (con la possibilità di ricerca delle parole nel testo), "<i>Imageviewer</i>", "<i>Photoviewer</i>" (per visualizzare l'immagine in diverse prospettive di grandezza), "<i>Videoviewer</i>" (Figura 22).</p>

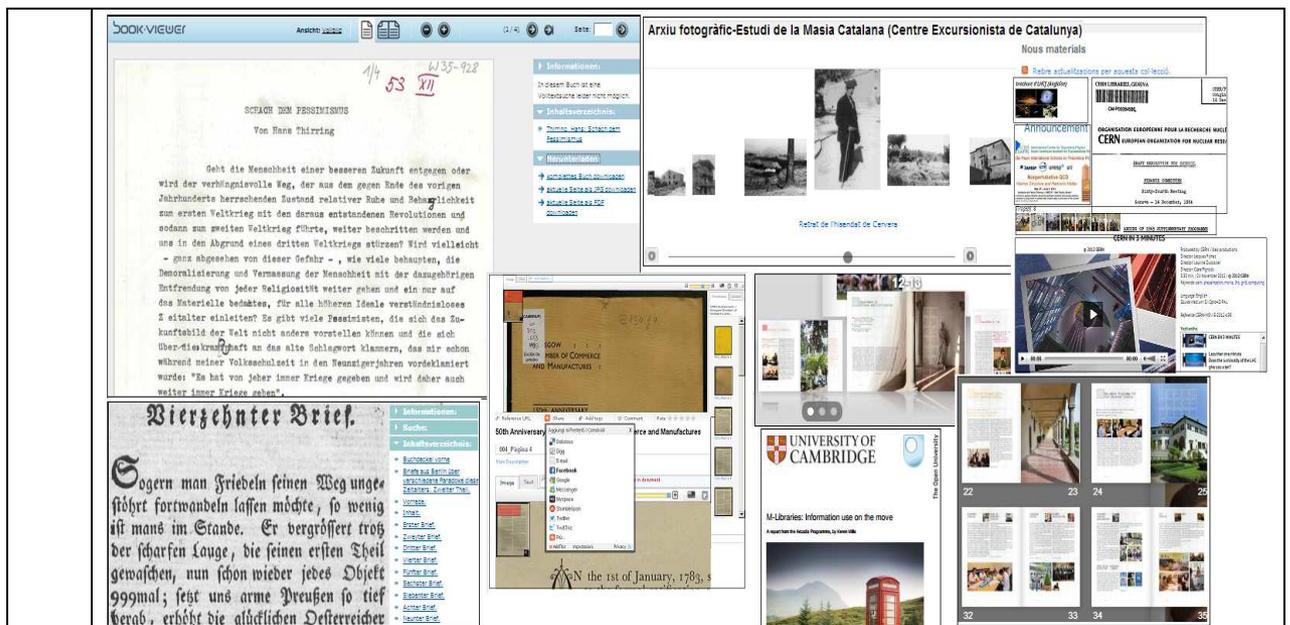


Figura 22. Alcune modalità di rappresentazione dei contenuti digitali nelle infrastrutture di ricerca oggetto d'indagine

Queste funzionalità sono pensate per un browsing rapido dei contenuti delle risorse e per catturare l'attenzione dell'utente.

4) Per arricchire le modalità di visualizzazione del contenuto dell'oggetto rilevato, è opportuno affiancarlo con i file allegati e/o dati di ricerca che possono riguardare lo stesso campo semantico e di ricerca.

Un oggetto digitale di ricerca affiancato dai file allegati prende il nome di "Pubblicazione potenziata" ovvero "Enhanced publication" [V.sito; Hogenaar, 2009].

I *Repository* indagati in cui è possibile visualizzare il contenuto dell'oggetto rilevato insieme ai file allegati sono:

- ✓ CERN Document Server;
- ✓ OpenAIRE Orphan Record Repository;
- ✓ Institutional Archive of Research AIR;
- ✓ Cambridge University DSpace;
- ✓ Publication Open Repository Torino (PORTO);
- ✓ PHAIDRA

5) È importante che l'oggetto depositato sia affiancato da funzionalità che permettono all'utente non solo di eseguire la sua visualizzazione, ma anche di poter eseguire le ricerche dentro il testo. Per questo è opportuno che:

a) i file PDF prevedano le possibilità di ricerca per parole chiave;

b) gli oggetti digitalizzati siano prodotti/scannerizzati secondo un riconoscimento ottico dei caratteri digitali, indipendentemente dalla sorgente delle immagini attraverso i programmi *OCR (Optical Character Recognition)*: OCR di caratteri stampati e di formattazione della pagina, OCR a mano libera, OCR del corsivo [V.sito: LAperLA; IA Bookreader].

6) Un buon esempio di *Repository* che permette all'utente di eseguire le ricerche entro i testi tramite le parole chiave, cliccando sui quali è possibile approfondire il contesto

semantico dell'oggetto di ricerca è fornito dall'*Europe PubMed Central Subject Repository*. La piattaforma di questo *Repository* incorpora le funzionalità della tecnologia *Text-mining*, con cui è possibile elaborare interi testi digitali rilevando al loro interno le parole con la frequenza maggiore. Le medesime parole vengono evidenziate e appoggiate sui link, attraverso i quali l'utente può consultare le informazioni correlate (es. dati di ricerca) contenute in diverse banche di dati con le quali l'*Europe PubMed Central* è collegato.

Anche l'*Oregon State University ScholarsArchive@OSU* incorpora le funzionalità di *Text-mining*, abbinando tale tecnologia con i meccanismi di *data linking* ai contenuti correlati per il campo semantico.

7) I *Repository* contenenti i testi digitalizzati di un certo valore storico-culturale non dovrebbero presentarli come “una semplice trasposizione in formato elettronico di edizioni cartacei” [V. “*Guida per un’edizione digitale dei manoscritti di Ferdinand de Saussure*”, p.2], bensì allestirli con opportuni strumenti professionali progettati per poter eseguire annotazioni e commenti critici dentro i medesimi testi.

La tecnologia informatica per la critica del testo “*Pinakes Text*” (*Strumento di filologia computazionale*) [V.sito] sviluppata dal CNR, Istituto di Linguistica Computazionale, Pisa (Italia) è uno tra tali strumenti (Figura 23):



Figura 23. La tecnologia informatica per la critica del testo “*Pinakes Text*” (*Strumento di filologia computazionale*)

Incorporando una simile tecnologia in un *Repository* contenete documenti digitalizzati antichi, il medesimo *Repository* acquisterà un notevole valore aggiunto, in quanto tra i potenziali utenti attirerà anche i critici del testo che possono lavorare, esaminare e facilmente annotare i contenuti dei medesimi documenti, accrescendo le loro meta-informazioni descrittive ed, eventualmente, rilevando anche il loro pregevole valore storico.

	7) Sarebbe auspicabile implementare in tutti i <i>Repository</i> il servizio “ <i>Printing-on-demand</i> ” per poter richiedere a distanza la possibilità di stampa dei contenuti visualizzati.
V. Comunicazione mediata dal <i>Repository</i> in collegamento con altri sistemi, moduli	
1.	Moduli di <i>Feedback</i>, <i>Forum</i>, <i>Blog</i> a supporto della comunicazione attiva con l’utenza
	<p>1) La maggior parte delle piattaforme è provvista di funzionalità di <i>Feedback</i>, mentre sono poche le piattaforme che offrono gli spazi per <i>Forum</i> e <i>Blog</i>. Esse sono:</p> <ul style="list-style-type: none"> ✓ Munich Personal RePEc Archive; ✓ Institutional Archive of Research (AIR); ✓ Europe PubMed Central Subject Repository; ✓ University College London ePrints; ✓ Research Archive “UCL Discovery”; ✓ The Virtual Open Access Agriculture & Aquaculture Repository VOA3R. <p>2) La questione sull’importanza dello sviluppo degli spazi <i>Blog</i> per la comunicazione dell’informazione scientifica è stata recentemente sollevata entro il progetto “<i>Hypothèses.org</i>” [V. sito], sfociato nel Portale per il <i>Blogg</i>ing in ambito accademico (<i>Scholarly blogging in the Social Sciences and Humanities</i>) disponibile in diverse lingue.</p> <p>3) I manager di <i>Repository</i> possono optare per l’implementazione di diverse tipologie di <i>Blog</i>:</p> <ul style="list-style-type: none"> ✓ <i>Blog written by a scientist,</i> ✓ <i>Blog written by a professional science writer/journalist,</i> ✓ <i>Blog that predominantly covers science topics,</i> ✓ <i>Blog used in a science classroom as a teaching tool,</i> ✓ <i>Blog used for more-or-less official news and press releases by scientific societies, institutes, centers, universities, publishers, companies and other organizations</i> [V. sito: “Science-at-risk-NDIIPP-report”].
	Moduli per la valutazione (<i>Rating</i>) dei contenuti
	La maggior parte dei <i>Repository</i> prevede gli spazi per il <i>Rating</i> dei contenuti, a partire dalle schede descrittive degli oggetti depositati
	Moduli per i commenti (<i>Comments</i>, <i>Discussions</i>) critici collaborativi sui contenuti
	È ‘opportuno che tutti i <i>Repository</i> siano allestiti con gli spazi – come “ <i>Comments</i> ”, “ <i>Discussions</i> ”- in cui gli utenti possono lasciare i loro commenti sui contenuti consultati.
	Applicazione delle tecnologie di <i>social tagging</i>
	La maggior parte dei <i>Repository</i> prevede possibilità di applicare le tecnologie di <i>social tagging</i> (“ <i>Share</i> ”, “ <i>Like</i> ”, “ <i>Comments</i> ”) con l’appoggio di vari canali di <i>Social Networks</i> , consentendo agli utenti accreditati di assegnare loro drescrittori e, pertanto, di aumentare il potenziale di aggregazione reticolare e associativa tra concetti, temi e oggetti.
	Collegamenti con i contenuti di altri sistemi
	<p>1) La maggioranza dei <i>Repository</i> indagati offrono i collegamenti/<i>link</i> per la consultazione dei contenuti correlati ai quelli depositati nei medesimi <i>Repository</i>. I collegamenti possono far riferimento ai contenuti distribuiti da altri <i>Repository</i>, ai cataloghi di riferimenti bibliografici online (interni ed esterni all’Istituzione), agli aggregatori online di diverse tipologie di contenuti.</p> <p>2) Bisognerebbe dotare ogni Archivio di ricerca di “una rete di link reciproci con le altre risorse e servizi informativi dell’Istituzione con cui l’Archivio è integrato [...] Servizi a</p>

	<p>valore aggiunto che possano renderlo più appetibile alla comunità accademica” [Guerrini, 2010, p.18].</p> <p>La creazione di link reciproci con le risorse informativi di sistemi esterni all’Istituzione potrebbe essere messa in opera a seconda degli accordi di collaborazione sviluppati tra le Istituzioni e altri soggetti produttori dei contenuti digitali.</p> <p>Per i <i>Repository</i> dei contenuti di ricerca e quelli di valore storico sarebbe, ad esempio, opportuno creare i collegamenti con le collezioni dei materiali aggregati:</p> <ol style="list-style-type: none"> a) nell’infrastruttura digitale dei materiali di ricerca “<i>OpenAIRE</i>” ; b) nel Portale dei contenuti digitali del patrimonio storico-culturale dell’Europa “<i>Europeana</i>”.
--	---

Tabella 10. Risultati dell’indagine e alcune raccomandazioni per potenziare le piattaforme di ricerca dei contenuti digitali

3.5. Conclusioni. I fattori catalizzanti il valore degli Archivi Istituzionali

Lo spazio di questo Capitolo è stato dedicato a descrivere un’indagine eseguita su 23 piattaforme (nazionali e internazionali) dei contenuti digitali a sostegno della comunicazione del sapere in rete e del Libero Accesso ai contenuti di ricerca. In questo studio è stato interessante osservare come le medesime piattaforme, volte inizialmente a realizzare qualche scopo ben preciso, siano diventate delle realtà concrete e abbastanza sviluppate, integrandosi nel tessuto nazionale ed internazionale degli strumenti e servizi digitali di ricerca.

Considerando i risultati della ricerca svolta, nelle presenti Conclusioni si cerca di riflettere sui maggiori fattori catalizzanti il valore e il successo degli Archivi Istituzionali dei materiali di ricerca (Tabella 11). I medesimi fattori possono essere presi in considerazione come alcuni degli elementi con cui allestire un eventuale *check-list* per valutare la qualità di programmazione dei contenuti negli Archivi Istituzionali.

<input type="checkbox"/>	<p>Ogni Archivio Istituzionale deve essere allestito con schema/i di metadati e loro mappatura/e ampiamente riconosciuto/i e condiviso/i a livello nazionale/europeo/internazionale, per favorire una migliore comunicazione e integrazione tra i contenuti di diversi Archivi e sistemi online.</p>
<input type="checkbox"/>	<p>Per migliore integrazione, comunicazione e interrogazione dei contenuti digitali, è opportuno appoggiare questi ultimi sui meccanismi di <i>Linked Data</i> [V. APPENDICE I].</p> <p>In questa direzione è opportuno imitare le esperienze di:</p> <ol style="list-style-type: none"> a) <i>servizi Linked Data per le Tesi e Dissertazioni dell’Oregon Institutional Repository</i> [Johnson, Boock, 2012]; b) <i>Repository VOA3R</i> e c) Portali “<i>The European Library</i>” e “<i>Europeana</i>” <p>sui quali server vengono implementati i sistemi di organizzazione della conoscenza KOS (Classificazioni, Thesauri) pubblicati online in modalità <i>Linked Data</i> e adoperati per qualificare i valori di metadati a supporto descrittivo dei contenuti;</p> <ol style="list-style-type: none"> d) creazione da parte di ZBW Labs [V.sito] dei metadati in modalità <i>Linked Data</i> per <i>Repository</i> disciplinare <i>Econstor</i> [V.sito]. <p>Ovviamente, ci sono tanti altri casi d’uso di piattaforme digitali i cui contenuti vengono affiancati dalla tecnologia <i>Linked Data</i> e per mezzo dei quali i medesimi contenuti entrano a fare la parte dello spazio dei <i>dataset</i> interconnessi del Web Semantico.</p> <p style="text-align: center;"><i>Complying with the Linked Data principles solves the problems of syntax and access</i></p>

	<p><i>heterogeneity by standardization. The problem of semantic heterogeneity remains but is eased in the Linked Data context as: Many Linked Data sources reuse terms from widely-used vocabularies (ontologies) to represent data about common types of entities such as people, products, reviews, publications, and other creative works. In addition, they use their own, proprietary terms for representing aspects that are not covered by the widely used vocabularies. [Bizer, 2013, p. 6].</i></p>
<input type="checkbox"/>	<p>È importante lo sviluppo di nuovi metodi per il <i>linking</i> al fine di esprimere le relazioni complesse tra gli oggetti entro un <i>Repository</i> e tra essi e i contenuti dello stesso campo semantico in altri <i>Repository</i>. Ad esempio, per poter facilmente leggere i contenuti correlati in un unico ambiente, possono essere <i>linkati</i>:</p> <ul style="list-style-type: none"> a) gli articoli ai dati di ricerca, al progetto di ricerca di riferimento; b) gli articoli alle parti di un manoscritto e/o lettere di corrispondenza digitalizzate, contenute in vari Archivi digitali. <p>In questa direzione sarebbe utile impiegare la già citata tecnologia <i>Linked Data</i>, per mezzo della quale le unità informative connesse possono arricchire la loro “espressività” concettuale, semantica e quella di contesto.</p> <p><i>Just as hyperlinks in the classic Web connect documents into a single global information space, Linked Data uses hyperlinks to connect disparate data into a single global data space. These links, in turn, enable applications to navigate the data space. For example, a Linked Data application that has looked up a URI and retrieved RDF data describing a scientific experiment may follow links from that data to data on different Web servers, describing related experiments [Bizer, 2013, pp. 2,4].</i></p> <p>Prendendo come base i collegamenti creati tra diverse unità informative è possibile sviluppare la visualizzazione dei grafi rappresentanti le parti di uno stesso contenuto e delle relazioni attraverso le quali poter seguire lo sviluppo dell'intero ciclo del progetto di ricerca, della scrittura del manoscritto e così via.</p>
<input type="checkbox"/>	<p>Dopo l'instaurazione dell'infrastruttura <i>OpenAIRE</i> per l'aggregazione dei contenuti di ricerca degli Archivi di ricerca <i>Open Access</i> nell'Unione Europea vi è la necessità di sviluppare i meccanismi per connettere gli Archivi di ricerca a livello globale, al fine di sviluppare un “<i>Global Scientific Data Space</i>”:</p> <p><i>The increasing global adoption of Linked Data technologies for sharing scientific data, library data and e-government data as well as the first generation of Linked Data discovery tools, such as Linked Data search engines, indicate the suitability of the Linked Data architecture for extending the Web with a Global Scientific Data Space [...] The Web of Linked Data is a realization of the Data Space vision on global scale [Bizer, 2013, p.5]</i></p>
<input type="checkbox"/>	<p>Sarebbe opportuno che ogni Archivio di contenuti di ricerca sia allestito con i collegamenti/<i>link</i> tra gli oggetti depositati entro e fuori l'Archivio.</p> <p>Il Portale NARCIS (<i>National Academic Research and Collaborations Information System</i>) [V. sito] fornisce un buon esempio per lo sviluppo delle funzionalità di ricerca diversificate che ritagliano uno spazio anche per la ricerca delle pubblicazioni potenziate, cioè “<i>Enhanced Publication</i>”: le pubblicazioni collegate ad una diversità di <i>dataset</i> prescelte (articoli, parti del libro, dati di ricerca ecc.) che appartengono allo stesso campo semantico.</p>
<input type="checkbox"/>	<p>L'implementazione di <i>identificatori unici persistenti, formati</i> a supporto della conservazione a lungo termine per ogni oggetto depositato - deve avere la priorità per assicurare i futuri accessi e la possibilità di consultazione del materiale a medio e lungo termine: “<i>long-lived digital data collections are powerful catalysts for progress and for democratization of science and education</i>” [V. sito: “<i>Long-Lived Digital Data Collections</i>”, p.5].</p>
<input type="checkbox"/>	<p>Ogni Università deve avere la consapevolezza di considerare il proprio Archivio Istituzionale come un utile strumento non solo per il deposito del materiale di ricerca,</p>

	<p>ma anche per la gestione delle questioni amministrative legate all'Anagrafe della ricerca, contribuendo così al monitoraggio della produzione scientifica ai fini della sua valutazione.</p> <p>In questa prospettiva gli Archivi Istituzionali possono diventare il fulcro di molte attività istituzionali e il loro uso deve entrare, perciò, nella pratica quotidiana di tutta la popolazione dei relativi Atenei:</p> <p>Gli Archivi Istituzionali stanno diventando un valido supporto per la valutazione della ricerca mettendo a disposizione nuovi dati che vanno ad affiancarsi ai quelli elaborati degli indici tradizionali e nello stesso tempo si presentano come una vetrina della produzione scientifica accademica consentendo forme di comparazione e <i>benchmarking</i> [V. sito: "Utilizzo degli Archivi Istituzionali per la valutazione della ricerca: esempi e opportunità"].</p>
<input type="checkbox"/>	<p>È importante lo sviluppo a monte di Archivi Istituzionali dei servizi a valore aggiunto come, ad esempio:</p> <p>a) funzioni "<i>Print on demand</i>";</p> <p>b) moduli per "Commenti critici collaborativi";</p> <p>c) moduli per "Annotazioni" sui contenuti digitali e digitalizzati;</p> <p>d) servizi per la comunicazione attiva con l'utenza come <i>Forum, Blog</i> ecc.</p>
<input type="checkbox"/>	<p>È opportuno implementare a monte degli Archivi delle piattaforme di pubblicazione come <i>Open Journal Systems</i> [V. sito] e di altri tipi di piattaforme <i>Open Source</i> per consentire ai Dipartimenti di impegnarsi in editoria <i>Open Access</i> e in altre attività innovative della "<i>Digital Scholarship</i>".</p>
<input type="checkbox"/>	<p>Per ogni Archivio Istituzionale è importante seguire un piano nazionale/europeo/internazionale per il prossimo futuro organico di sviluppo e di gestione degli Archivi, come, ad esempio, proposto dalle Linee Guida CRUI, <i>DRIVER</i>, <i>OpenAIRE</i> [V.siti].</p>
<input type="checkbox"/>	<p>Dopo l'instaurazione dell'infrastruttura <i>OpenAIRE</i> per l'aggregazione dei contenuti di ricerca degli Archivi di ricerca <i>Open Access</i> nell'Unione Europea vi è la necessità di sviluppare i meccanismi per connettere gli Archivi di ricerca a livello globale, al fine di sviluppare un "<i>Global Scientific Data Space</i>":</p> <p><i>The increasing global adoption of Linked Data technologies for sharing scientific data, library data and e-government data as well as the first generation of Linked Data discovery tools, such as Linked Data search engines, indicate the suitability of the Linked Data architecture for extending the Web with a Global Scientific Data Space [...] The Web of Linked Data is a realization of the Data Space vision on global scale</i> [Bizer, 2013, p.5]</p>

Tabella 11. Maggiori fattori che possono catalizzare il valore e il successo di Archivi Istituzionali

Per dare pieno avvio all'attuazione di indirizzi e criteri [V. sito: "*Criteria for successful Repositories*"] che possano catalizzare il valore di Archivi Istituzionali, è importante:

- 1) il supporto di Archivi Istituzionali da parte degli Organi di Governo delle Università;
- 2) la dotazione da parte delle Università di Politiche interne ben definite a favore della promozione del materiale di ricerca tramite gli Archivi Istituzionali implementati;
- 3) il maggiore coinvolgimento degli autori nelle attività di deposito dei propri lavori di ricerca nei medesimi Archivi;
- 4) un'attiva promozione dell'iniziativa di Archivi Istituzionali, che è un motore per la continuità, la crescita e la cooperazione dei progetti iniziati.

Considerando un continuo progresso nel campo ICT (*Information Technology and Information*) e una progressiva globalizzazione dell'informazione, ogni Archivio Istituzionale deve aggiornare e potenziare in continuazione le proprie funzionalità relativamente alla presentazione, ricerca e

disseminazione della conoscenza sul materiale di ricerca. Quest'ultimo, a sua volta, deve rafforzare il proprio profilo verso l'apertura alla comunicazione della conoscenza scientifica, ancora di più se la produzione di tale conoscenza è supportata dai fondi pubblici. Ciò deve avvenire tramite la più ampia adozione delle strategie e politiche che si stanno promuovendo tramite il movimento internazionale *Open Access*, cui scopo è quello di aumentare la disseminazione dei contenuti di ricerca su larga scala tramite la rete. Consolidando la cooperazione tra diverse istituzioni produttori dei contenuti digitali, una simile prospettiva per lo sviluppo di Archivi digitali di ricerca non farà l'altro che giovare al progresso della ricerca stessa a livello globale.

È opportuno che i fattori che possono catalizzare il valore di Archivi Istituzionali siano formalizzati entro le Politiche di Archivi Istituzionali. Inoltre, sarebbe auspicabile che ogni *Repository* sia allestito con un *Toolkit* che raggruppa sotto le opportune categorie le informazioni relative a risorse, aspetti teorici e Buone Pratiche ritenute importanti nel campo applicativo di sviluppo e di gestione degli Archivi Istituzionali, incluso la pianificazione delle funzionalità e degli strumenti a supporto dei contenuti depositati.

Nel prossimo Capitolo è presentato un *Toolkit* cui contenuto è strutturato in categorie contenenti brevi descrizioni e riferimenti per l'approfondimento delle informazioni relative a supporto dell'implementazione e della gestione di Archivi digitali di ricerca.

CAPITOLO 4

ToolKit basato sullo stato dell'arte di alcuni aspetti a supporto del "Profilo Tecnico" di Archivi digitali di ricerca

4.0. Introduzione

Con l'avvento di Internet l'offerta di canali informativi per la disseminazione dei materiali digitali che possano presentare un interesse per la ricerca scientifica sta crescendo. Questi canali spaziano da quelli per la ricerca dell'informazione ad uso comune (es. *Google; Yahoo; Bing*) a quelli opportunamente creati per disseminare il materiale intellettuale prodotto nell'ambito accademico (es. *Google Scholar*; portale web per la ricerca di Tesi di Dottorato, articoli, monografie, rapporti scientifici - *Scholarius; Scirus*), a quelli sviluppati e/o ospitati dalle Biblioteche accademiche come, ad esempio:

- ✓ OPAC SBN (Catalogo del Servizio Bibliotecario Nazionale),
- ✓ JSTOR (Biblioteca digitale di riviste accademiche, libri e fonti scientifiche primarie),
- ✓ *Scopus* (database di riassunti e citazioni per articoli di pubblicazioni riguardanti la ricerca ad alto *impact factor*; consente di calcolare *H-index*),
- ✓ DOAJ (*Directory of Open Access Journals*);

In aggiunta a ciò, ormai da decenni i ricercatori in tutto il mondo auto-archiviano e disseminano la conoscenza sui loro lavori di ricerca tramite gli Archivi aperti (Istituzionali e/o Disciplinari) *Open Access* indicizzabili e ricercabili tramite i canali web come:

- ✓ DOAR (*Directory of Open Access Repositories*);
- ✓ ROAR (*Registry of Open Access Repositories*);
- ✓ *Web Ranking of World Repositories*;
- ✓ OAISTER (*Union Catalog of digital resources*).

Inoltre, diversi contenuti di Archivi *Open Access* vengono aggregati e indicizzati per la ricerca da mega-piattaforme come:

- ✓ NARCIS (*National Academic Research and Collaborations Information System*);
- ✓ *Networked Digital Library of Theses and Dissertations* (NDLTD);
- ✓ DART-Europe *E-theses Portal*;
- ✓ DRIVER (*Digital Repository Infrastructure Vision for European Research*);
- ✓ RePEc (*Repository of Research Papers in Economics*);
- ✓ VOA3R (*Virtual Open Access Agriculture & Aquaculture Repository*);
- ✓ *Europeana* (una mega-biblioteca digitale europea);
- ✓ WorldCat (un mega-catalogo online delle collezioni di biblioteche a livello mondiale);
- ✓ OpenAire (*Open Access Infrastructure for Research in Europe*) [V.siti].

Di buona norma e dal punto di vista tecnico-funzionale, la progettazione e la realizzazione delle infrastrutture dedicate all'allestimento e alla diffusione delle risorse informativi digitali di valore scientifico-culturale devono:

1. ottemperare ai requisiti specifici di *best practice*, modelli, specifiche e standard internazionali e come obiettivi primari

2. garantire la massima interoperabilità con analoghi sistemi alla base di progetti nazionali ed internazionali e allo stesso tempo supportare tipologie di risorse digitali e metadati eterogenei, dai più "tradizionali" derivanti da progetti di digitalizzazione di collezioni bibliografiche ai dataset prodotti da simulazioni scientifiche e altre attività di ricerca.

Questa prospettiva potrebbe rilevarsi molto importante per un miglioramento della posizione di un canale per la disseminazione dei contenuti digitali "tra i pari". Ad esempio, analizzando periodicamente il servizio del *Web Ranking of World Repositories* si potrebbe osservare come diversi Archivi "guadagnano" diverse posizioni "tra i pari" a seconda dei criteri con cui viene

valutata la loro capacità nella gestione e disseminazione dei contenuti [V. sito: “*Repositories webometrics*”].

L’obiettivo del *Toolkit* fornito nel presente Capitolo è quello di allestire alcuni elementi del “Profilo Tecnico” [V. sito: “Linee Guida per gli Archivi Istituzionali” 2009, p. 12] di un Archivio digitale di ricerca, riportando sinteticamente lo stato dell’arte di una serie di componenti a supporto funzionale dei medesimi Archivi. Questi componenti - schematizzati nella Figura 24 - saranno descritti con le informazioni introduttive e affiancati da riferimenti di approfondimento (informazioni relative ad alcuni progetti, metodologie, attività, servizi e iniziative pratiche maturate nel campo a livello nazionale e internazionale). In questo modo il lettore sarà introdotto alla conoscenza di diversi elementi specifici e comuni che possono essere applicati nel contesto degli Archivi Istituzionali.

Sarebbe senz’altro proficuo che l’interfaccia di ogni Archivio Istituzionale sia allestita con una sezione dedicata alla descrizione di una serie di aspetti metodologici e tecnico-funzionali aggregati in un *ToolKit*, al fine di:

- ✓ favorire l’orientamento dell’utente “non esperto” e promuovere la conoscenza degli esperti relativamente allo stato dell’arte di alcuni modelli, strumenti e servizi a supporto dell’architettura orientata alla gestione, ricerca e rappresentazione dei contenuti digitali negli Archivi;
- ✓ fornire un continuo attivo confronto tra il livello metodologico e quello applicativo nel campo di riferimento;
- ✓ fornire un set di strumenti di valutazione oggettiva e misurabile, tale da consentire la creazione di una *check-list* durante una predisposizione di elementi di sintesi dello stato dell’arte degli aspetti per la cura della programmazione, gestione e manutenzione dell’Archivio;
- ✓ poter passare all’applicazione pratica di alcuni aspetti ricavati dallo stato dell’arte al fine di maturare nuovi servizi operativi e un miglior modello dell’Archivio.

Per di più, sarebbe opportuno che un *ToolKit* integrato nella piattaforma di un Archivio sia potenziato dai canali digitali di comunicazione online attiva come *Feed-back*, *Blog* - attraverso cui gli utenti registrati possano continuamente fornire loro commenti e/o proposte per valutare, migliorare e ottimizzare le performance dell’Archivio in questione.

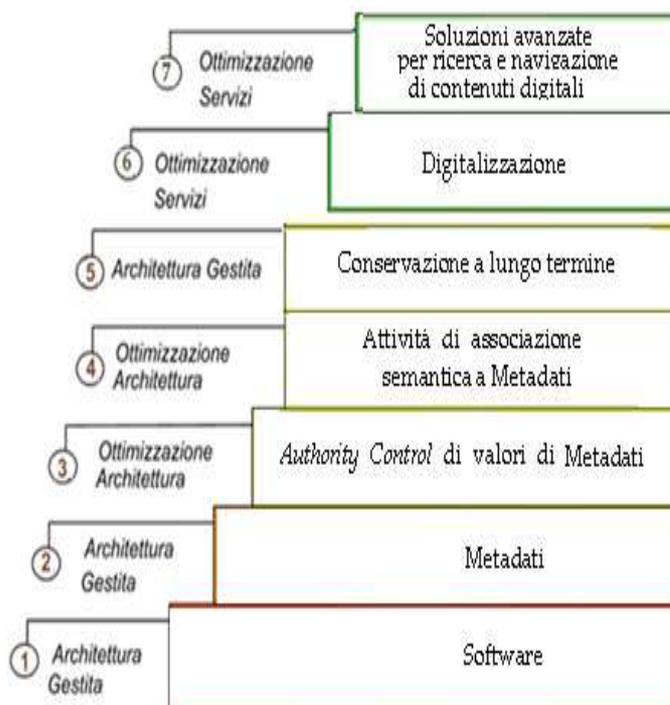


Figura 24. Alcuni componenti del “Profilo Tecnico” di un Archivio digitale di ricerca

Un buon esempio di come una piattaforma digitale può incorporare al proprio interno una sezione informativa dedicata alle questioni relative alla gestione dei contenuti digitali viene presentata dalla piattaforma online AIMS (*Agricultural Information Management Standards*) [V.sito], la pagina principale di cui ritaglia uno spazio per all’allestimento dei seguenti aspetti:

- ✓ “Open Access”,
- ✓ “Vocabolari controllati”,
- ✓ “Metadati”,
- ✓ “Tools”,
- ✓ “Linked Data” (Figura 25) .

Attraverso queste e altre sezioni la piattaforma AIMS connette la conoscenza di specialisti da tutto il mondo, che intervengono con propri inserzioni (*banner*) per contribuire con le loro informazioni a una comunità di pratica interconnessa a livello globale.



Figura 22. Piattaforma online AIMS (*Agricultural Information Management Standards*)

Il *Toolkit* che verrà presentato nel presente Capitolo può essere considerato un "condensato" delle informazioni che descrivono *good practice* e le linee progettuali teoriche e pratiche (Figura 24) di un impianto tecnologico dell’Archivio digitale di ricerca, con una particolare attenzione agli aspetti di integrazione tra le medesime linee e l’ambiente di un Archivio aperto Istituzionale.

I contenuti con cui è stato allestito il *Toolkit* non hanno la pretesa di essere esaustivi. Possono essere piuttosto i punti da cui partire per elaborare un’ulteriore strategia per il potenziamento del “Profilo Tecnico” degli Archivi Istituzionali, attraverso i passi da eseguire, in maniera incrementale per il raggiungimento di livelli di sviluppo successivi o, in maniera di riadattamento alle esigenze di funzionalità aggiuntive ed identità funzionali richieste dal progetto dinamico di sviluppo delle funzionalità di un Archivio Istituzionale.

4.1. Panorama delle piattaforme

Ciò che serve a livello di cooperazione italiana è una *piattaforma* per l'accesso agli strumenti *Software OAI*, che sia di supporto alla costruzione di Archivi per la ricerca, ma anche per la didattica, Istituzionali o Disciplinari, alla creazione di servizi per la raccolta e il *caching*¹ di metadati, *harvester*, per il *reference linking* entro gli *Open Archives*, per conversione di metadati da un formato all'altro, per l'esportazione e importazione di metadati, per la ricerca nel testo pieno dei documenti [De Robbio, 2007, pp. p. 258-259].

In seguito saranno esposte informazioni, elencate fonti e forniti rimandi per approfondire il raggio del campo teorico e operativo relativo al nodo centrale di tutte le infrastrutture digitali – le piattaforme *Software*.

“Linee Guida per gli Archivi Istituzionali” (CRUI, Italia):

<<http://www.cru.it/HomePage.aspx?ref=1781>>

I *Software Open Source* più diffusi nel mondo che offrono le fondamenta architettoniche e lo strato di orchestrazione dei processi negli Archivi aperti sono:

- ✓ DSpace: <www.dspace.org>;
- ✓ EPrints: <www.eprints.org>;
- ✓ Fedora (*Flexible Extensible Digital Object Repository*): <<http://www.fedora.info/>>;
- ✓ CDS *Invenio* (originariamente CDSware): <<http://cdsware.cern.ch/invenio/index.html>>.

I *Software Open Source* possono essere scaricati e aggiornati gratuitamente, il codice sorgente può essere liberamente esaminato, personalizzato e migliorato, così come il controllo di sicurezza.

Per implementare un Archivio Istituzionale si potrebbe optare per l'installazione di:

- ✓ *Software Open Source* “in casa” (*in house*, locale);
- ✓ *Software* esternalizzato (*hosted*, servizi in *outsourcing*);
- ✓ eventuali soluzioni ibride.

Spesso le soluzioni locali contrapposte a quelle “*hosted*” sembrano essere meno costose, ma occorre valutare quali siano le esigenze del servizio, che spesso solo un *Data Center* professionale può soddisfare.

Caratteristiche necessarie di un *Software*:

- ✓ affidabilità elevata e *Business Continuity*;
- ✓ connettività; sicurezza informatica;
- ✓ sicurezza logistica e fisica;
- ✓ monitoraggio e gestione;
- ✓ *backup* (salvataggi centralizzati automatizzati, salvaguardia Basi Dati);
- ✓ *storage*;
- ✓ *disaster recovery site*.

Il *Software* scelto per l'installazione dell'Archivio Aperto Istituzionale deve supportare lo Standard di interoperabilità tecnica OAI-PMH (*Open Archives Initiative – Protocol for Metadata Harvesting*) 2.0: <<http://www.openarchives.org/pmh/>>,

la prima e principale caratteristica tecnica dell'applicazione scelta per gli Archivi aperti (*Open Archives*) che riguardano *Data* e *Service Provider*.

OAI-PMH permette:

- ✓ l'interoperabilità minima tra i contenuti di Archivi aperti appoggiandosi sui metadati semplici *Dublin Core* [V. APPENDICE A].
- ✓ *Service Provider* di attingere tramite i metadati raccolti i contenuti da *Data Provider*

¹ *Cache* è una raccolta di dati che rappresenta la duplicazione di valori originali memorizzati altrove su un computer.

opportunamente integrati: <<http://www.openarchives.org/Register/BrowseSites>>. In questo modo *Service Provider* creano i servizi centralizzati a valore aggiunto come è offerto dai diversi Portali e *Directory*;

✓ la ricerca di contenuti di diversi Archivi aperti tramite i motori di ricerca generalisti e specialistici (*Google, Yahoo, Google Scholar, Oaister, ecc.*).

□ *Software* scelti per l'implementazione di Archivi aperti Istituzionali devono essere in grado di interagire tramite le interfacce Standard e Web Service con altri sistemi locali ed esterni dell'ente, per poter offrire servizi quali:

✓ l'autenticazione (autorizzazione) centralizzata via *LDAP, CAS, Shibboleth*;

✓ la cattura (raccolta) di informazioni da database dell'Ateneo (es. l'Anagrafica del personale, le descrizioni bibliografiche dal Portale di Biblioteca);

✓ la fornitura di informazioni a database (es. la lista delle pubblicazioni per il rapporto annuale dell'ente, per il Sito Docente del Ministero).

Analisi comparativa di 11 *Software* (commerciali e *Open*). Progetto “*Repositories Support Project*” (UK), 2010: <<http://www.rsp.ac.uk/start/software-survey/results-2010>>

□ La comparazione di *Software* è stata eseguita in base ad un set di caratteristiche definite a monte. Per ogni *Software* sono stati forniti gli esempi di *Repository* implementati (Figura 26, a).

CONTENTdm		Digital Commons		DigiTool	DSpace	EPrints
Commercial	Commercial	Commercial	Commercial	Free	Free	Free
Software	Hosted service	Software	Software	Software	Software	Software

QUELLA Repository	Fedora	Islandora/Fedora	intraLibrary	Open Repository	Zentity
Commercial	Free	Free	Free	Commercial	Free
Software	Software	Software	Software	Hosted service	Software

Figura 26 (a). Analisi comparativa di 11 *Software* (commerciali e *Open*) eseguita entro il progetto “*Repositories Support Project*”

□ L'indagine del Progetto ha rilevato le seguenti piattaforme *Software* per l'implementazione di Archivi digitali aperti:

✓ *Software proprietario (commerciale)*. Per la sua implementazione si paga una tariffa e, opzionalmente, ogni ulteriore consulenza o abbonamento per ottenere *hosting, backup offsite*, supporto tecnico, aggiornamenti. Con una programmazione dell'interfaccia API (*Application Programming Interface*) è possibile personalizzare il *Software*, ma il codice sorgente rimane in possesso dal fornitore del *Software*;

✓ *Software Open Source/Free (download gratuito)*. Il codice sorgente può essere liberamente esaminato, personalizzato e migliorato;

✓ i fornitori di *Hosted services* come “*Open Repository*” e “*Digital Commons*” si prestano a gestire i dati di acquirenti dei medesimi servizi, fornendo funzionalità aggiuntive, controllando e aggiornando il codice sorgente.

Il prospetto della mappatura di alcuni principali aspetti in base a cui è stata eseguita l'analisi comparativa è il seguente (Figura 26, b):

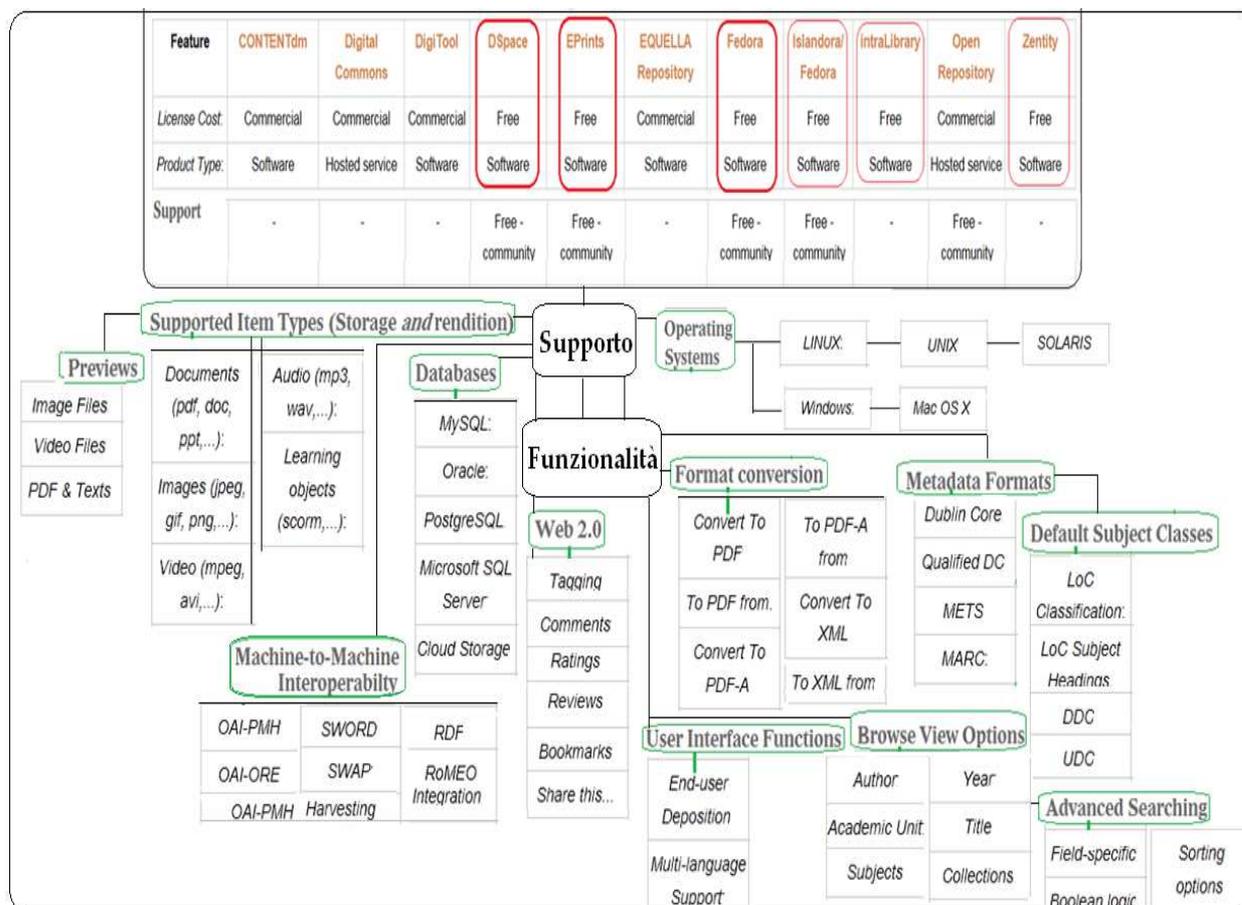


Figura 26 (b). Mappatura di alcuni principali aspetti in base a cui è stata eseguita l'analisi comparativa dei Software

□ Sotto vengono fornite le descrizioni succinte dell'analisi dei Software Open Source DSpace, EPrints, Islandora/Fedora.

Dspace	EPrints	Islandora/Fedora (basato su Drupal ² e Fedora)
< http://www.dspace.org >; < https://wiki.duraspace.org/display/DSPACE/RoadMap >	< http://www.eprints.org/software/ >	< http://islandora.ca/ >; < http://www.fedora-commons.org/about/examples/islandora/ >
> Supporto multi-lingua	Supporto multi-lingua	Supporto multi-lingua
Repositories (Archivi) implementati e indicizzati dalla Directory OpenDOAR		
Dspace	EPrints	Islandora/Fedora
< http://www.opendoar.org/find.ind.php?rSoftwareName=D >	< http://www.opendoar.org/find.php?rSoftwareName=EPr >	< http://www.opendoar.org/find.php?rSoftwareName=Islandora >

² DRUPAL e' un Content Management System (CMS) Open Source molto diffuso. Ha l'architettura modulare, flessibile e scalabile che permette di assemblare e personalizzare vere e proprie applicazioni web (es. visualizzazione di risorse e risultati di ricerca; gestione autorizzazioni, accessi a collezioni e risorse all'interno del Repository; monitoraggio dell'ingestione e transcoding; realizzazione di browsing selettivi su alcune porzioni di collezioni) integrate con servizi esterni.

[Space&format=titles&step=50](http://www.opendoar.org/Space&format=titles&step=50) > [ints&format=titles&step=50](http://www.opendoar.org/ints&format=titles&step=50) > [%2F+Fedora&format=titles&step=50](http://www.opendoar.org/%2F+Fedora&format=titles&step=50)

Repositories (Archivi) implementati e indicizzati dal servizio web Worldwide Google Map

Dpace	EPrints	Islandora/Fedora
http://www.opendoar.org/find.ind.php?rSoftWareName=DSpace&format=gmap	http://www.opendoar.org/find.ind.php?rSoftWareName=EPrints&format=gmap	http://www.opendoar.org/find.php?rSoftWareName=Islandora%2F+Fedora&format=gmap

Sistemi operativi per l'implementazione dei Software

Dpace	EPrints	Islandora/Fedora
LINUX, UNIX, SOLARIS, Windows, MacOS	LINUX, UNIX, SOLARIS, Windows, MacOS	LINUX, UNIX, SOLARIS, Windows, MacOS, Hosted service, Cloud Server

Interoperabilità con altri sistemi tramite protocolli:

- ✓ OAI-PMH (*Open Archives Initiative – Protocol for Metadata Harvesting*);
- ✓ OAI-ORE (*Open Archives Initiative Object Reuse and Exchange*);
- ✓ SWORD (*Simple Web-service Offering Repository Deposit*);
- ✓ WebDAV (*Web Distributed Authoring and Versioning*);
- ✓ REST (*Representational State Transfer*).

Integrazione con modelli, sistemi, servizi:

- ✓ SWAP (*Scholarly Work Application Profile*),
- ✓ RDF *Resource Description Framework*,
- ✓ SHERPA/RoMEO (*Publisher copyright policies & self-archiving*);
- ✓ SRU/SRW (*Search and Retrieve Web Service and Search and Retrieve URL Service*).

Dpace	EPrints	Islandora/Fedora
OAI-PMH, OAI-ORE, SWORD, SWAP, WebDAV, REST, SRU/SRW	OAI-PMH, OAI-ORE, SWORD, SWAP, RDF, SHERPA/RoMEO <i>Integration</i> (attraverso componenti aggiuntivi da fornitori esterni)	OAI-PMH, OAI-ORE, SWAP (opzionale), SWORD (attraverso componenti aggiuntivi da fornitori esterni), RDF, <i>RoMEO Integration</i> , SRU/SRW

Software che supportano il *data model* RDF possono integrare la tecnologia del Web Semantico *Linked Data*, adatta per una programmazione consistente dei collegamenti tra *dataset*³ identificati con *http_URI*.

Con l'appoggio della tecnologia *Linked Data* i *dataset* del *Repository* saranno abilitati per la loro esposizione e l'eventuale integrazione (creazione di collegamenti in base all'affinità semantica e contesto) in un ecosistema di *dataset* e database interconnessi in rapida crescita dello spazio virtuale del Web Semantico - *Linked Open Data (LOD) Cloud* (<http://linkeddata.org/>).

Nel modello grafico del *LOD Cloud* vengono rappresentati diversi *dataset*, abilitati, pubblicati e interconnessi come *Linked (Open) Data* dai vari soggetti produttori (biblioteche, archivi, musei, pubbliche amministrazioni, aziende pubbliche e private ecc.), arrecando così un altissimo valore aggiunto ad una diversità di *dataset* pubblicati sul web.

Tipologie di contenuti e formati supportati

Dpace	EPrints	Islandora/Fedora
Documenti (pdf, doc, ppt,...).	Documenti (pdf, doc, ppt,...). Immagini (jpeg, gif, png,...).	Supporta l'installazione di tutte le tipologie.

³ "A *dataset* is a collection of data, published and maintained by a single Provider", <https://wiki.duraspace.org/display/DSPACE/Linked+Open+Data+for+DSpace>

Immagini (jpeg, gif, png,...).	Video (mpeg, avi,...).
Video (mpeg, avi,...).	Audio (mp3, wav,...).
Audio (mp3, wav,...).	Materiale didattico (scorm,...).
Materiale didattico (scorm,...).	Componenti aggiuntivi da fornitori esterni.

Conversione e anteprima dei formati

Dpace	EPrints	Islandora/Fedora
Per convertire file in PDF/PDF-A, XML - installare componenti aggiuntivi da fornitori esterni. Per l'anteprima di <i>Image Files; Video Files, PDF & Texts</i> - installare componenti aggiuntivi da fornitori esterni.	È prevista la conversione in PDF (caratteristiche fanno parte del pacchetto, ma devono essere attivate). È possibile l'anteprima di <i>Image Files; Video Files, PDF & Texts</i> .	È prevista la conversione in <i>PDF, in PDF-A, in XML da: od*, ot*, sxw, doc, docx, txt, htm, pdb, xls, csv, xml, wpd, wb2, pp*, mml</i> . È possibile l'anteprima di <i>Image Files; Video Files, PDF & Texts</i> .

Database supportati

Dpace	EPrints	Islandora/Fedora
<i>Oracle, PostgreSQL</i>	MySQL, Oracle, PostgreSQL, <i>Cloud Storage</i>	MySQL, Oracle, PostgreSQL, <i>Cloud Storage</i>

Vocabolari controllati e Classificazioni per soggetto supportati

Dpace	EPrints	Islandora/Fedora
Supporto di qualsiasi vocabolario controllato definito dall'amministratore del sistema.	<i>LoC (Library of Congress Classification)</i> ; altri vocabolari controllati.	Supporto di qualsiasi vocabolario controllato.

Funzionalità di "Browsing" per le voci:

"Autore", "Comunità"/"Unità istituzionale", "Soggetto", "Anno", "Titolo", "Collezioni"

Dpace	EPrints	Islandora/Fedora
"Autore", "Soggetto", "Anno", "Titolo", "Collezioni". <i>Configurabile per ogni altro campo di navigazione.</i>	"Autore", "Unità istituzionale", "Soggetto", "Anno". <i>Configurabile per ogni altro campo di navigazione.</i>	"Autore", "Unità istituzionale", "Soggetto", "Anno", "Titolo", "Collezioni". <i>Configurabile per ogni altro campo di navigazione.</i>

Funzionalità di "Ricerca avanzata"

Dpace	EPrints	Islandora/Fedora
Specificazione dei campi. Logica Booleana. Opzioni di ordinamento	Specificazione dei campi. Opzioni di ordinamento	Specificazione dei campi. Logica Booleana. Opzioni di ordinamento

Funzionalità del Web 2.0

(*tagging, commenti, valutazioni, revisioni, segnalibri, condivisione*)

Dpace	EPrints	Islandora/Fedora
Condivisione: componenti aggiuntivi dai fornitori esterni.	<i>Tagging</i> , commenti: componenti aggiuntivi da fornitori esterni.	<i>Tagging</i> , commenti, valutazioni, revisioni, segnalibri, condivisione. Altre funzionalità via <i>Drupal</i> [Corlosquet et al., 2011].

“Confronto fra Software per la gestione di *Digital Repository*”. Progetto Magazzini Digitali: <<http://www.rinascimento-digitale.it/documenti/confrontorepository/ConfrontoRepository2006.pdf>>

Nel presente progetto sono stati esaminati i *Software DSpace, Fedora* (Figura 27), *Digitool, aDORe, Greenstore, MyCoRe* secondo i seguenti criteri di inclusione:

- supporto per lo Standard OAIS (*Open Archival Information System*) per la conservazione degli oggetti digitali a lungo termine. Il modello internazionale a pacchetti OAIS definisce AIU (*Archival Information Unit*) per quanto riguarda gli oggetti e AIC (*Archival Information Collection*) per quanto riguarda le collezioni;
- possibilità di autenticazione tramite LDAP (*Lightweight Direct Access Protocol*) o SSO (*Single Sign On*), fondamentali per l'integrazione dei *Software* con i sistemi di autenticazione locali;
- supporto per gli identificatori *persistenti URL*, importanti nell'ottica della persistenza dell'accesso dei contenuti digitali a lungo termine;
- capacità di gestire file di tipo più eterogeneo possibile;
- possibilità di interazione con altri sistemi: integrazione con Portali, con gli ambienti VLE (*Virtual Learning Environments*), con i sistemi di gestione di biblioteche digitali;
- supporto per il protocollo OAI-PMH, essenziale per rendere disponibili online i metadati alle ricerche dei contenuti;
- content versioning*;
- supporto dello Standard ANSI Z39.50, essenziale per l'integrazione delle piattaforme con i servizi di Biblioteca digitale, e la sua evoluzione in SRU/SRW;
- supporto per l'*OpenURL*, un protocollo per lo scambio di metadati finalizzato alla gestione di servizi di *linking* contestualizzato (*context-sensitive*) per l'integrazione con i servizi offerti dalla Biblioteca digitale.

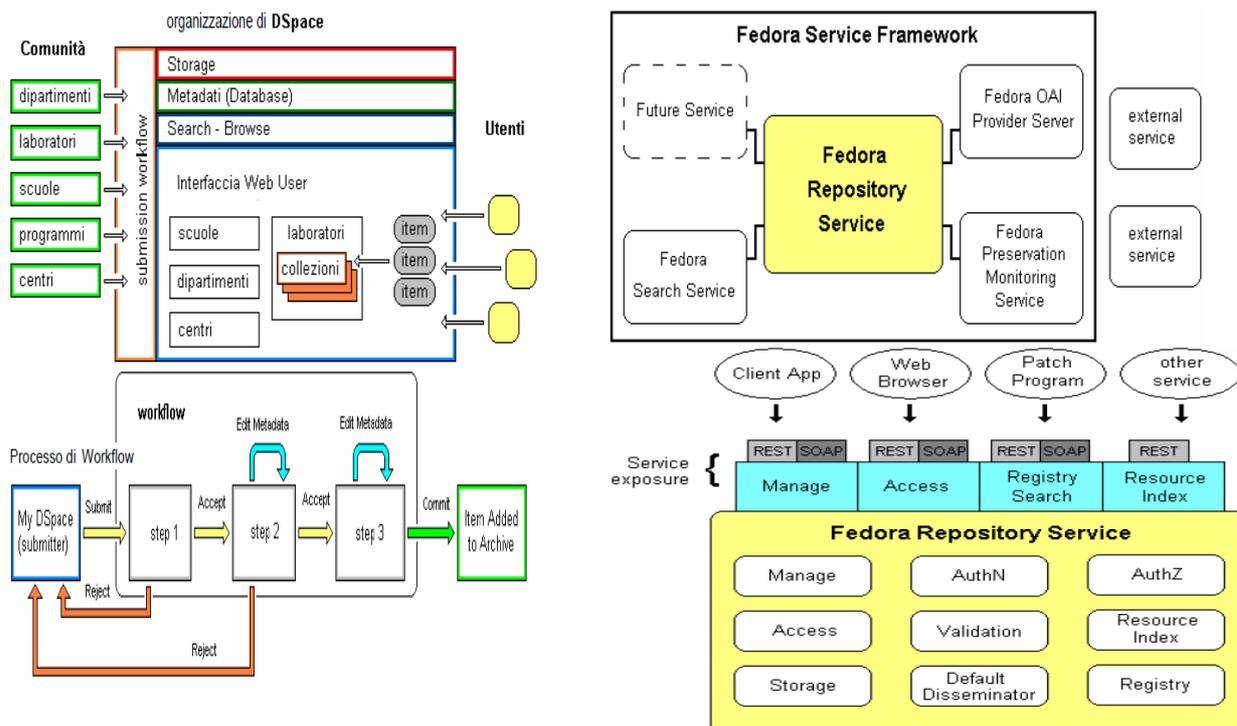


Figura 27. Rappresentazione grafica dell'architettura dei *Software DSpace* e *Fedora*

Software Open Source Eprints. Versioni 3.2.0; 3.2.1:

<http://wiki.eprints.org/w/New_Features_in_EPrints_3.2>;

<<http://blogs.ukoln.ac.uk/ap-support/2010/03/23/linked-data-and-dublin-core-application-profiles-in-eprints-3-2-0/>>

Le recenti versioni del Software EPrints prevedono le funzionalità avanzate come il supporto per profili applicativi⁴ di metadati sviluppati sul *Dublin Core Application Profile (DCAP)* appoggiato sui costrutti della c.d. “*FRBR Family*”:

- ✓ FRBR (*Functional Requirements for Bibliographic Records*),
- ✓ FRAD (*Functional Requirements for Authority Data*),
- ✓ FRSAD (*Functional Requirements for Subject Authority Data*).

Il primo ad usare esplicitamente il modello FRBR attraverso DCAP è stato il profilo applicativo SWAP (*Scholarly Works Application Profile, DC-SWAP model*):

<http://www.ukoln.ac.uk/repositories/digirep/index/Eprints_Application_Profile>

sviluppato per formalizzare gli elementi rappresentativi del materiale di ricerca e l'interazione dei medesimi espressa attraverso i collegamenti entro un modello.

Altri Software per l'implementazione di Archivi digitali

□ **Archimede Open Source Software:** <<http://www1.bibl.ulaval.ca/archimede/index.en.html>>; <<http://sourceforge.net/projects/archimede>>

- ✓ viene distribuito sotto la licenza generale pubblica GNU;
- ✓ offre una prospettiva multilingue (lingue di base: Inglese, Francese, Spagnolo);
- ✓ la sua interfaccia è indipendente e non è incorporata nel codice. Questo permette: di sviluppare ulteriori interfacce di lingue specifiche senza la re-codifica del *Software*

⁴ Un *Profilo Applicativo* è uno schema di metadati elaborato da uno o più *namespace* di metadati che, messi insieme secondo certe regole, possono ottimizzare una particolare applicazione locale del sistema informativo digitale, <<http://files.eprints.org/471/>>

stesso; agli utenti di passare da una lingua all'altra durante la ricerca e il recupero di contenuti;

✓ l'esempio di installazione: *Laval University Library*: <<http://www.bibl.ulaval.ca/>>

□ **Bepress commercial Software**: <<http://www.bepress.com/repositories.html>>:

✓ è il partner di *ProQuest Information&Learning (Digital Commons@)* e le sue prestazioni vengono armonizzate con quelle della biblioteca digitale *ProQuest* per il deposito delle tesi e dissertazioni (ETDs);

✓ l'esempio di installazione: *University of California's eScholarship Repository*:

<<http://repositories.cdlib.org/escholarship/>>

□ **Greenstone: Software Libero multi-lingue Open source**:

<<http://www.greenstone.org/>>; <<http://www.greenstone.org/download>>:

✓ è distribuito sotto la licenza pubblica generale GNU;

✓ offre una prospettiva multilingue (lingue di base: Inglese, Francese, Spagnolo, Russo);

✓ l'esempio di installazione: *Oxford Digital Library*:

<<http://www2.odl.ox.ac.uk/gsd/cgi-bin/library/>>

□ **Rosetta (di Ex Libris) commercial Software**:

<<http://www.exlibrisgroup.com/category/RosettaOverview>>

fornisce un sistema altamente scalabile, sicuro e facile per gestire e preservare contenuti digitali.

□ **Software Selection Methodology for Library Discovery Layer Systems**:

<<http://foss4lib.org/book/export/html/4779>>

□ **Software per navigare ed esplorare i contenuti FOSS Discovery Layer**:

✓ **VuFind**: <<http://www.foss4lib.org/package/vufind>>;

✓ **Blacklight**: <<http://www.foss4lib.org/package/blacklight>>

offrono una forte alternativa ai Software commerciali come *Aquabrowser*, *Primo* o *Encore*.

Fedora Digital Object Model:

<<https://wiki.duraspace.org/display/FEDORA34/Fedora+Digital+Object+Model>>

□ L'architettura del *Software Fedora* definisce un modello di *Oggetto Digitale* generico che può essere utilizzato per fornire le caratteristiche essenziali per molti tipi di oggetti digitali, incluso documenti, immagini, libri elettronici, oggetti di apprendimento multimediali, set di dati, metadati e molti altri.

□ Questo modello è un elemento fondamentale per l'architettura di *Content Model* e per tutte le altre funzionalità previsti da *Fedora*, particolarmente per quanto riguarda rendere *persistenti* molteplici tipologie di oggetti digitali. *Content Model* è un modello formale che descrive le caratteristiche e i servizi dei contenuti digitali (es. "*cModel:aggregation*"; "*cModel:genericVideo*"; "*cModel:videoSource*"; "*cModel:collection*"). Riduce lo sforzo di cattura, acquisizione, archiviazione, gestione, preservazione, validazione, trasformazione e accesso dei contenuti digitali, fornendo anche un metodo di classificazione degli stessi.

□ Attraverso la *Content Model Architecture* i contenuti digitali *non sono definiti in base al formato dei file o in base alla tecnologia* e incorporano i servizi che agiscono sui contenuti stessi.

DSpace 2.0 Data Model:

<<https://wiki.duraspace.org/display/DSPACE/DSpace+2.0+Expressing+DSpace+Domain+Model+In+RDF>>

DSpace 2.0 Data Model esprime un modello di dati che include metadati e ogni relazione che potrebbe essere creata tra diversi oggetti del *Repository*, attraverso il linguaggio del Web Semantico RDF (*Resource Description Framework*).

Servizi per la governance dei contenuti di ricerca di Atenei: U-GOV-Ricerca, SURplus -

Open Archive

L'integrazione di tali servizi con l'Archivio Istituzionale rientrano nelle attività di supporto allo standard (protocollo) di interoperabilità OAI-PMH e offre la possibilità di realizzare Archivi aperti sul web facilitando la reperibilità e la fruibilità dei risultati di ricerca degli Atenei su motori di ricerca nazionali e internazionali.

U-GOV-Ricerca: <<http://www.cineca.it/page/u-gov-e-ricerca>>;
http://www.cineca.it/sites/default/files/Cineca_DigitalUniversity.pdf

- ✓ È una suite informatizzata per la gestione e il monitoraggio delle attività di ricerca a livello di Ateneo;
- ✓ offre le basi per razionalizzare l'utilizzo delle risorse, ottimizzare la gestione dei progetti, verificare il raggiungimento degli obiettivi, valutare i risultati e le competenze acquisite;
- ✓ l'Archivio Istituzionale in collegamento (tramite OAI-PMH) con U-GOV-Ricerca raccoglie i metadati esposti da quest'ultimo, li memorizza e indicizza per la ricerca entro l'Archivio. L'Archivio sta a valle di un processo di validazione e autorizzazione dei metadati da U-GOV-Ricerca;
- ✓ tramite le funzionalità di *harvesting* OAI-PMH è possibile anche definire una collezione dell'Archivio alimentata con i Prodotti della Ricerca di uno specifico dipartimento dell'Ateneo.

SURplus - Open Archive: <<http://www.cilea.it/index.php?id=surplus>> è un "Sistema Università Ricerca plus - Open Archive".

I vantaggi offerti dall'integrazione dei servizi per la *governance* con gli Archivi Istituzionali:

- ✓ l'apertura controllata dei metadati verso sistemi esterni;
- ✓ l'interazione dei contenuti degli Archivi con i dati dei sistemi interni amministrativi (es. Anagrafe, Banche Dati di dipartimenti ecc.) ed esterni ai relativi Atenei (es. banche dati di MIUR);
- ✓ i ricercatori e quindi gli Atenei hanno la possibilità di valorizzare i Prodotti della Ricerca indicizzati dai servizi la *governance*, rendendoli ricercabili e maggiormente visibili tramite gli Archivi Istituzionali.

Ghent Declaration:

<http://www.openaire.eu/index.php?option=com_content&view=article&id=223%3Aseizing-the-opportunity-for-open-access-to-european-research-ghent-declaration-published&catid=76%3Ahighlights&lang=it>

Questa Dichiarazione - sollecitata dalla Commissione Europea - invita gli stati dell'Unione Europea a cogliere l'opportunità fornita dal Progetto OpenAIRE per migliorare e incentivare la circolazione online della conoscenza scientifica tramite le piattaforme compatibili con l'infrastruttura OpenAIRE [V. sito: "Compatibilità delle piattaforme *DSpace*, *EPrints* e quelle proprietarie con l'OpenAIRE"].

Dalla scelta del Software all'implementazione dell'Archivio Istituzionale

Una volta si è deciso di implementare un Archivio Istituzionale, è necessario eseguire un esame comparativo dei requisiti di Software a supporto di tale compito disponibili sul mercato [Fay, 2010].

In un secondo momento è opportuno restringere il campo della selezione secondo i criteri definiti per il soddisfacimento dei bisogni informativi locali del sistema e dell'utenza.

La scelta del *Software* dipenderà anche dalle risorse disponibili, considerando i costi complessivi per: *Software*, *Hardware*, *Server*, *staff* operativo, *staff* responsabile per

programmazione delle funzionalità aggiuntive, *backup*, *disaster recovery*⁵, recupero, preservazione a lungo termine dei contenuti digitali.

La scelta di una certa piattaforma *Software* richiede la presenza di diverse competenze professionali aventi l'*expertise* di come un sistema deve operare.

Una volta fatta la scelta della piattaforma, si passa all'assemblaggio e alla configurazione di *Software*, *Hardware*, e *Server* necessari.

Dopo l'installazione di una versione *demo* e al *test* del sistema si passa alla realizzazione di un *Repository Pilota*.

V. RUBRIC Toolkit: “*Establishing a Pilot Repository*”; “*Transition from the Pilot system to a production system*”: <http://docsfiles.com/pdf_rubric_toolkit_pilot_to_production.html>

È opportuno programmare le *funzionalità personalizzate* dell'Archivio (es. presentazione interfaccia; ricerca e navigazione tramite le voci controllate).

Formazione dello *staff* a supporto delle funzionalità e dei flussi dell'Archivio programmati.

Indipendentemente dalla soluzione del *Software* per un Archivio Istituzionale (IR) scelta, è importante avere consapevolezza del fatto che *limitarsi* ad una installazione base *Software* può comportare una *perdita di impatto* dell'Archivio. Le *personalizzazioni*, ad esempio, delle funzionalità dell'interfaccia che possono essere applicate a livello di sistema per potenziare la ricerca e navigazione delle informazioni

su misura per i propri utenti è un *punto chiave* per il *successo* di un Archivio Istituzionale e può contribuire ad attrarre ulteriore considerazione e finanziamenti.

È pur vero d'altra parte che, nel disegno di personalizzazione, va accuratamente accertata la fattibilità: la *personalizzazione “spinta”* di un Archivio che utilizza un *Software Open Source* potrebbe comportare un eccessivo dispendio di risorse umane e finanziarie per l'allineamento del sistema e *problemi nella migrazione* da una versione all'altra [...]

In caso di *personalizzazioni* delle *funzionalità* e/o dei *processi*, si raccomanda di *documentare le modifiche* e le ragioni per non perderne traccia nel tempo, in particolare in occasione delle *migrazioni*

[V. sito: “Linee Guida per gli Archivi Istituzionali”, pp. 12, 26]

“È necessario che le *installazioni* esistenti siano sempre *costantemente aggiornate*, per poter sfruttare al meglio le grandi potenzialità dei *Software* e le sempre nuove funzionalità richieste“ [Guerrini, 2010a, p. 50].

⁵ Con l'entrata in vigore del D.Lgs. 235/2010 diventano operative le modifiche al CAD predisposte dal Ministero per la Pubblica Amministrazione e l'Innovazione: Diventa OBBLIGATORIA la definizione di un PIANO di CONTINUITÀ OPERATIVA e di un PIANO di *DISASTER RECOVERY* (le Misure Tecniche e organizzative per garantire il funzionamento dei centri di elaborazione dati e delle procedure informatiche rilevanti in siti alternativi a quelli di produzione. Le soluzioni tecniche idonee a garantire la salvaguardia dei dati e delle applicazioni informatiche, verifica annualmente il costante aggiornamento dei piani di *DISASTER RECOVERY* delle amministrazioni interessate e ne informa annualmente il Ministro per la pubblica amministrazione e l'innovazione).

4.2. Metadati a supporto dei contenuti digitali

L'attività di metadattazione è finalizzata all'individuazione di set di metadati (tra gli Standard internazionali e nazionali esistenti) più idonei a descrivere le collezioni e le singole risorse documentarie. I metadati:

- ✓ sono la colla che tiene insieme le informazioni sugli oggetti digitali, garantendone la qualità e l'accesso, e, al contempo,
- ✓ rappresentano la struttura logica che consente flussi (fruizione) di informazioni e nuove relazioni tra le risorse digitali;
- ✓ favoriscono l'identificazione univoca, l'individuazione e la localizzazione di una risorsa;
- ✓ facilitano l'aggregazione e l'organizzazione dinamiche di risorse con caratteristiche comuni;
- ✓ possono garantire la conservazione a lungo termine delle rappresentazioni digitali delle risorse documentarie.

Metadata plays a crucial role in the functioning of individual Repositories, and also networks of Open Access Repositories [...] Because of the crucial role that metadata plays in making the content of Repositories discoverable and searchable not only within an individual Repository, but also within a network of them, its creation is an extremely important step of the submission process
[Gunning, p.3].

Per ogni oggetto e collezione si applicherà una metadattazione leggera (di base, es. *Dublin Core simple*) dal punto di vista descrittivo e conservativo.

Il profilo descrittivo dei materiali dell'Archivio Istituzionale "non è un Catalogo, è uno strumento fondamentalmente diverso e quindi sia l'interfaccia sia i metadati devono essere "a portata di mano" dell'utente finale, così come lo schema disciplinare per il *browsing* utile alla collocazione entro uno schema disciplinare o anche puramente gerarchico che dovrà essere il più semplice possibile
[De Robbio, 2007, p.106].

La qualità di metadati si raggiunge tramite la conformità agli Standard e l'arricchimento delle proprietà di metadati con elementi semantici (soggetti, temi, generi ecc.), generati utilizzando Schemi di Classificazione e vocabolari controllati (*Authority File*, ontologie, Thesauri), a seconda del dominio disciplinare di appartenenza dei singoli oggetti e delle aggregazioni per collezioni.

Il rapporto tra la qualità di metadati e gli Archivi Istituzionali "riveste un'importanza primaria considerando che la qualità dei primi ha una profonda incidenza sul livello qualitativo dei servizi offerti all'utente.

La qualità dei metadati riflette il grado con cui i metadati esercitano le fondamentali funzioni bibliografiche di ricerca, uso, diffusione, autenticità e gestione in un contesto di interoperabilità, quale quello promosso dal modello OAI, la qualità dei metadati utilizzati nei *Data Provider* influenza sia i servizi agli utenti locali, sia i servizi forniti dai *Service Provider* e può determinare gravi problemi nel trasferimento dati nel caso di migrazioni tra piattaforme diverse (anche compatibili con le specifiche del protocollo OAI-PMH).

La ricerca di risorse in rete risulta molto più agevole se si usano Schemi di metadati definiti, protocolli di trasferimento condivisi e mappature tra gli schemi.

[Guerrini, 2010 a, p. 42].

In seguito saranno espresse informazioni, elencate fonti, e forniti rimandi di approfondimento relativamente al tema del presente paragrafo.

"Understanding metadata" (NISO):

<<http://www.niso.org/publications/press/UnderstandingMetadata.pdf>>

L'introduzione al mondo di metadati (registri, schemi, standard).

"Linee Guida per gli Archivi Istituzionali" (CRUI, Italia):

<<http://www.cruil.it/HomePage.aspx?ref=1781>>

Scegliere metadati che possono adeguatamente coprire gli aspetti *descrittivi, strutturali, tecnici, amministrativi, di diritti* per rappresentare e gestire le risorse informative depositate nell'Archivio.

Procedere con una adeguata mappatura tra diversi schemi di metadati.

Per una maggiore visibilità dell'Archivio Istituzionale, anche a livello europeo, i metadati dovranno essere compatibili con le specifiche del progetto DRIVER⁶, particolarmente con le Linee Guida “*Digital Repository Infrastructure Vision for European Research*” 2.0: <<https://issue.guidelines.driver.research-infrastructures.eu/>>, sulla base di cui sono state sviluppate le Linee Guida “OpenAIRE” 1.1:

<<http://www.openaire.eu/component/content/article/207>>.

Le Linee Guida DRIVER possono servire ai gestori di:

- ✓ Archivi Istituzionali nuovi per definire le Politiche relative ai formati delle risorse e ai metadati (mappatura di metadati nel formato *Dublin Core*);
- ✓ Archivi già avviati per poter, ove è necessario, riorientare le scelte fatte.

Se un Archivio corrisponde alle caratteristiche ritenute indispensabili dalle Linee Guida DRIVER, esso acquista lo status di “*DRIVER Provider*” e diventa il partner dei *Repository* accolti entro l'infrastruttura digitale di DRIVER. Il primo requisito obbligatorio per la compatibilità con il progetto DRIVER è la presenza entro gli Archivi Istituzionali delle risorse informative a testo pieno (*full-text*).

Metadati Dublin Core: <<http://www.dublincore.org/documents/usageguide/#whatismetadata>>

V. ALLEGATO A

Di buona norma, alla base della descrizione delle risorse digitali negli Archivi Istituzionali devono essere i metadati *Dublin Core* (DC) semplici (15 elementi) - *Dublin Core Metadata Element Set (DCMES)*. Tramite questi elementi – che nell'ambiente web vengono supportati dal protocollo dell'interoperabilità tecnica OAI-PMH - gli Archivi (*Data e Service Provider*) possono comunicare tra loro.

Nonostante gli elementi di metadati DC possono essere ripetibili, al fine di creare più descrittori per rappresentare una risorsa, l'adozione del solo DC semplice non è in grado di coprire tutte le esigenze di rappresentazione dei contenuti digitali in un Archivio di ricerca.

Dublin Core Metadata Element Set (DCMES) nella versione italiana:

<http://www.iccu.sbn.it/opencms/opencms/it/main/standard/metadati/pagina_116.html>,

pubblicato dall'Istituto Centrale per il Catalogo Unico delle biblioteche italiane e per le informazioni (ICCU).

1. **Titolo** (un nome dato alla risorsa);
2. **Creatore** (un'entità che ha la responsabilità principale della produzione del contenuto della risorsa);
3. **Soggetto** (l'argomento della risorsa);
4. **Descrizione** (una spiegazione del contenuto della risorsa);
5. **Editore** (un'entità responsabile della produzione della risorsa, disponibile nella sua forma presente);
6. **Autore di contributo subordinato** (un'entità responsabile della produzione di un contributo al contenuto della risorsa);
7. **Data** (una data associata a un evento del ciclo di vita della risorsa).
8. **Tipo** (la natura o il genere del contenuto della risorsa);

⁶ DRIVER è un progetto di un Consorzio Europeo (formato da enti di ricerca, università, biblioteche, progetti nazionali) finanziato dalla Unione Europea, che offre servizi avanzati atti a raccogliere, aggregare ed organizzare le risorse presenti nei diversi Archivi Istituzionali e Disciplinari.

- 9. **Formato** (la manifestazione fisica o digitale della risorsa);
- 10. **Identificatore** (un riferimento univoco alla risorsa nell'ambito di un dato contesto);
- 10. **Fonte** (un riferimento a una risorsa dalla quale è derivata la risorsa in oggetto);
- 11. **Lingua** (la lingua del contenuto intellettuale della risorsa);
- 12. **Relazione** (un riferimento alla risorsa correlata);
- 13. **Copertura** (l'estensione o scopo del contenuto della risorsa);
- 14. **Gestione dei diritti** (l'informazione sui diritti esercitati sulla risorsa).

“Open Metadata Handbook/Metadata Standards”:

<http://en.wikibooks.org/wiki/Open_Metadata_Handbook/Metadata_Standards#BibJSON>

Modelli di dati, schemi di metadati, protocolli di metadati

Metadati “item-level”. Mappature (crosswalks) tra elementi di diversi schemi di metadati

□ Nella prassi descrittiva delle risorse digitali depositati negli Archivi Istituzionali, il set di metadati DC semplice (elementi generali) rappresenta spesso la base per un più ricco corredo di metadati “item-level” (Figura 28) il quale di buona norma deve essere documentato attraverso le mappature (crosswalks) tra diversi elementi di metadati scelte per raffinare il livello della granularità della descrizione.

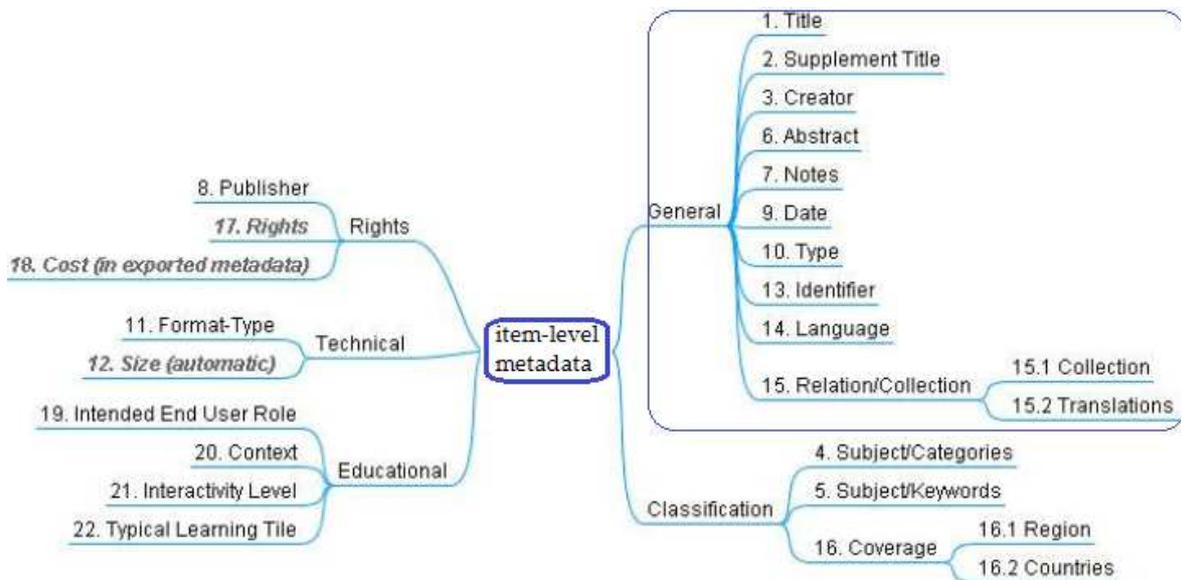


Figura 28. Esempio di uno schema concettuale di metadati “item-level”

Esempi di metadati “item-level”

□ Profili applicativi e Schemi di metadati:

✓ *Dublin Core Libraries Working Group Application Profile (LAP):*

<<http://dublincore.org/documents/library-application-profile/>>.

Dallo schema LAP possono essere attinti dei qualificatori per gli elementi DC. Altri qualificatori possono essere creati in autonomia (es. valori di metadati opportunamente qualificati attraverso le voci di vocabolari controllati).

✓ *Dublin Core Application Profiles (DCAP):*

<<http://dublincore.org/documents/profile-guidelines/index.shtml>>

✓ *Images Application Profile (IAP):*

<http://www.ukoln.ac.uk/repositories/digirep/index/Images_Application_Profile>

✓ *Geospatial Application Profile (GAP):*

<http://www.ukoln.ac.uk/repositories/digirep/index/Geospatial_Application_Profile>

- ✓ *Learning Materials Application Profile (DC Education AP) (LMAP)*:
<<http://www.icbl.hw.ac.uk/lmap/lmapscopingreport.pdf>>
- ✓ *Time-Based Media Application Profile (TBMAP)*:
<<http://www.jisc.ac.uk/whatwedo/programmes/reppres/tbmap.aspx>>
- ✓ *Scholarly Works Application Profile (SWAP)*:
<http://www.ukoln.ac.uk/repositories/digirep/index/Eprints_Application_Profile>
- ✓ *Metadata Object Description Schema (MODS)*: <<http://www.loc.gov/standards/mods/>>.

I Software più diffusi per la creazione di archivi OA (per esempio *Eprints* e *DSpace*) prevedono Profili Applicativi di metadati estensibili che, in fase di esportazione, possono essere codificati in diversi formati e schemi come MODS, RDF [Guerrini, 2010 a, p.42]

- ✓ *Bibliographic Ontology (BIBO)*: <<http://bibliontology.com/>>
- ✓ *Learning Object Metadata (IEEE LOM)*: <<http://ltsc.ieee.org/wg12/>>
- ✓ *Metadata for Learning Resources (MLR)*, uno schema compatibile con i metadati DC e IEEE LOM: <http://en.wikipedia.org/wiki/ISO/IEC_19788>
- ✓ XMetaDiss (metadati specifici per la descrizione delle Tesi):
<http://www.nb.admin.ch/nb_professionnel/01693/01699/01873/01894/index.html?lang=it&download=NHZLpZeg7t,lnp6I0NTU042I2Z6lnIah2oZn4Z2qZpnO2Yuq2Z6gpJCDeHx2hGym162epYbg2c_JjKbNoKSn6A-->>

Per approfondire l'argomento su *crosswalks* si vedono le seguenti fonti:

- ✓ “*Linee guida tecniche per i programmi di creazione di contenuti culturali digitali*”: Edizione italiana 2.0. Progetto Minevra, il Gruppo di lavoro italiano "Interoperabilità e servizi" WP4, 2007:
<<http://www.minervaeurope.org/publications/lineeguidatecnicheita/riperimento.html>>
- ✓ *Metadata Standards Crosswalk*:
<http://www.getty.edu/research/conducting_research/standards/intrometadata/crosswalks.html>
- ✓ *MARC Mappings*: <<http://www.loc.gov/marc/marcdocz.html>>
- ✓ *MODS Mappings*: <<http://www.loc.gov/standards/mods/mods-conversions.html>>
- ✓ Mappatura tra LOM e DC [Koutsomitropoulos, 2010]
- ✓ Mappatura tra DC e UNIMARC:
<http://www.iccu.sbn.it/opencms/opencms/it/main/standard/metadati/pagina_118.html>
- ✓ Mappatura tra DC e RDF: <<http://dublincore.org/documents/dc-rdf/>>

Applicazione delle mappature (crosswalks) tra elementi di diversi schemi di metadati comporta i vantaggi ai Repository in termini di interoperabilità interna ed esterna di loro contenuti.

□ L'uso delle mappature ampiamente condivise tra *Data Provider* (DP, Archivi digitali) e *Service Provider* (SP, servizi federati, aggregatori di Archivi) farà sì che i metadati sviluppati nei DP non saranno appiattiti a 15 elementi di metadati *DC simple* durante il processo di loro raccolta (*harvesting*) da parte di SP attraverso il protocollo OAI-PMH, bensì affronteranno coerentemente tutte le conversioni necessarie (Figura 29).

□ I metadati raccolti da DP vengono interpretati da parte di SP con il loro successivo confezionamento in diversi tipi di informazione per fornire servizi a valore aggiunto

[De Robbio, 2012, p. 12].

□ Esempi di SP:

✓ ArXiv: <<http://arxiv.org/>>

È un *Repository* su modello accentrato di *e-print* (materiale digitale) dei seguenti campi: fisica, matematica, scienza non lineare, informatica, biologia quantitativa e statistica. Offre la possibilità di meta-ricerca; è possibile quindi effettuare un'unica ricerca simultanea in queste ed altre banche dati, attraverso la ricerca multipla.

✓ RePEc: <<http://repec.org/>>

È un *Repository* federato che contiene la più ampia collezione distribuita al mondo di articoli accademici sull'economia e sulle discipline collegate. RePeC collega e raccoglie informazioni sui lavori, pubblicati e non, di migliaia di economisti che hanno depositato i loro lavori nei *Repository* locali che aderiscono al Network.

✓ SSRN:

<<http://papers.ssrn.com/sol3/DisplayJournalBrowse.cfm#>>;

<<http://repositories.webometrics.info/toprep.asp>>

È un *Repository* su modello accentrato a sé che "indicizza" il materiale nelle Scienze Umane solo dei *Repository* aderiti. Offre i servizi basati su specializzazioni disciplinari. Contiene anche papers ad accesso chiuso che vengono venduti. Nell'edizione di luglio 2011 di *Ranking web of world Repositories*, SSRN è stato considerato il miglior *Repository* dell'anno.

✓ DART-Europe (E-Thesis Portal): <<http://www.dart-europe.eu/basic-search.php>>;

✓ NARCIS (*National Academic Research and Collaborations Information System*):

<<http://www.narcis.nl/>>;

✓ OaiSter: <<http://www.oclc.org/oaister/>>;

✓ DRIVER (*Digital Repository Infrastructure Vision for European Research*):

<<http://search.driver.research-infrastructures.eu/Help.action?type=collection>>;

<<http://www.driver-repository.eu/>>

Specifiche di integrazione e specificazione di metadati (documenti del Progetto VOA3R)

□ “*Specifications for the integration of components*”:

<<http://agroknow.redpanda.gr/content/d421uahdraft20110221-specifications-integration-components>>

Le soluzioni di integrazione tecnica per ciascuno dei componenti del Sistema (*Repository*) comprendono la configurazione dei componenti per poter lavorare con i protocolli aperti per l'integrazione e la configurazione di metadati, per adattarli ai Profili di applicazione e *mapping*.

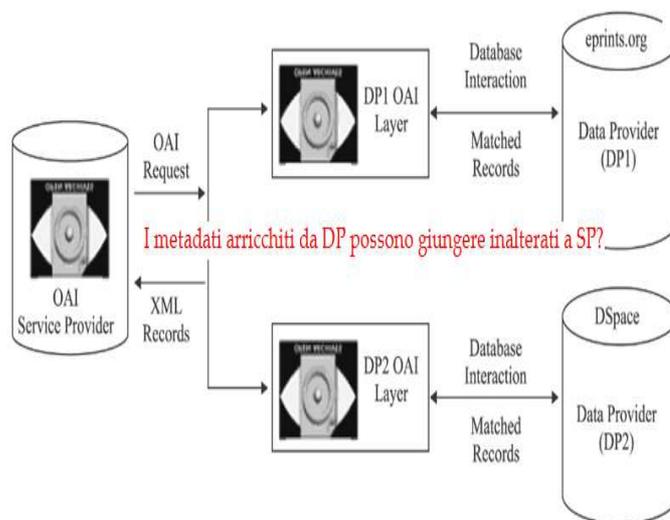


Figura 29. Schema di interazione tra *Data* e *Service Provider* OAI

□ “Specification of Metadata Profiles and mappings to existing technology”:

<<http://www.ieru.org/voa3r/wiki/images/b/bb/20120913-D3.5-PartB.pdf>>

“Resource Description and Access” (RDA). Standard compatibile con lo schema di metadati Dublin Core: <<http://www.ppt2txt.com/r/ffc282c8/>>

RDA può essere applicato per il controllo della forma di descrizione bibliografica.

Elenco di alcuni importanti “Metadata Value Standards and Resources”

<http://www.neal-schuman.com/metadata-digital-collections/MDC_Resources.html>

“Repositories Support Project” (UK): Software e schemi di metadati

<<http://www.rsp.ac.uk/>>

Il progetto specifica gli schemi di metadati supportati nativamente da diversi *Software* (Tabella 12).

Metadata Formats

Feature	CONTENTdm	Digital Commons	DigiTool	DSpace	EPrints	EQUELLA Repository	Fedora	Islandora/Fedora	intraLibrary	Open Repository	Zenity
Dublin Core:	Dublin Core	Dublin Core	Dublin Core	Dublin Core	Dublin Core	Dublin Core	Dublin Core	Dublin Core	Dublin Core	Dublin Core	Dublin Core
Qualified DC:	Qualified DC	Qualified DC	Qualified DC	Qualified DC	-	Qualified DC	Qualified DC	Qualified DC	Qualified DC	Qualified DC	-
METS:	METS	-	METS	METS	METS	METS	METS	METS	-	-	METS
MARC:	-	-	MARC	-	-	MARC	-	MARC	-	-	-
Other:	-	-	MODS, EAD, TEI, Can hold any metadata format	DSpace can import/export content from a variety of metadata formats, including MODS, PREMIS, etc.	BibTeX, DIDL, EndNote, JSON, MODS, CSV, ORE, OpenURL, RDF, Refer, RefMan, XML	can hold any Metadata format incl. SCORM and LOM IEEE (flexible)	Any XML metadata format	Any metadata format (e.g. VRACore, PBCore, mzXML, TEI)	LOM, XCRI, MIX, ODRL	Can import/export content from a variety of metadata formats, including MODS, MARC, PREMIS, etc)	Others can be defined by customer (including relationships)

Tabella 12. Schemi di metadati supportati da diversi *Software*

Supporto del formato di metadati semantici RDF (*Resource Description Framework*)

Le recenti versioni dei *Software DSpace* e *Fedora* supportano il formato dati RDF per esprimere le relazioni tra le triple (“Soggetto”/Risorsa - ha - “Predicato”/Proprietà - ha - “Oggetto”/Valore).

V. i documenti:

✓ **DSpace 2.0 - Expressing DSpace Domain Model in RDF:**

<<https://wiki.duraspace.org/display/DSPACE/DSpace+2.0+Expressing+DSpace+Domain+Model+In+RDF>>

✓ **Fedora Commons. All Documentation:**

<<https://wiki.duraspace.org/display/FEDORA/All+Documentation>>

Attraverso il modello dati RDF - diversi *dataset* (metadati, unità concettuali predefinite, voci dei vocabolari controllati, altri oggetti digitali) - identificati univocamente da *http_URIs* possono essere pubblicati in modalità *Linked (Open) Data* sul Web Semantico (Web di dati, Web 3.0) di dati interconnessi.

V. il documento “*Linked Open Data for DSpace*”:

<<https://wiki.duraspace.org/display/DSPACE/Linked+Open+Data+for+DSpace>>

Oggetti digitali rappresentati tramite i meccanismi *Linked (Open) Data* diventano non solo “*machine readable*” ma anche “*machine linkabile*”, quindi possono essere collegati e riusati da diversi agenti (produttori di *Linked Data* a loro volta) operanti sul Web di dati, quindi anche al

di fuori dell'originale contesto di utilizzo. Tale approccio offre una straordinaria possibilità a diversi di essere interconnessi tra di loro in base all'affinità semantica o all'area di interesse, indipendentemente dai formati e dai *Software* utilizzati per la loro produzione, liberando, allo stesso tempo, le istituzioni culturali dalla "dipendenza" da un limitato numero di fornitori tecnologici specializzati.

V. il documento "**Library Linked Data Incubator Group Final Report**". *W3C Incubator Group Report 25 October 2011*, <<http://www.w3.org/2005/Incubator/lld/XGR-lld-20111025/>>

"Metadata Handbook": <http://en.wikibooks.org/wiki/Open_Metadata_Handbook>

È un'iniziativa lanciata dal *Public Domain* e dall'*Open Bibliographic Working Group* dell'OKF (*Open Knowledge Foundation*) che sta mettendo in pratica le Raccomandazioni del Gruppo *Library Linked Data*.

"OpenAIRE Guidelines":

<<http://www.openaire.eu/en/component/content/article/9-news-events/427-openaire-releases-version-20-of-the-openaire-guidelines>>

Sono le specifiche per i metadati degli Archivi aperti *Open Access* pubblicate sulla base delle Linee Guida DRIVER:

OpenAIRE Guidelines 1.1 (*Guidelines for content providers of the OpenAIRE information space*): <<http://www.openaire.eu/component/content/article/207>>;

OpenAIRE Guidelines 2. 0 <<http://www.openaire.eu/en/component/content/article/9-news-events/427-openaire-releases-version-20-of-the-openaire-guidelines>>.

Attraverso *OpenAIRE Guidelines* manager di Repository possono definire Politiche locali per la gestione di metadati in conformità con le esigenze dell'Area Europea di Ricerca (ERA), con l'obiettivo di contribuire alla creazione di una infrastruttura digitale capace di *raccogliere, arricchire e conservare* i metadati dei set di dati scientifici e per *sostenere* la *gestione* e il *collegamento* tra i dati scientifici ad Accesso Aperto.

Secondo "*OpenAIRE Guidelines*" la descrizione di una risorsa di ricerca digitale entro un Archivio ad Accesso Aperto deve essere obbligatoriamente affiancata dai seguenti metadati:

- ✓ *Title*,
- ✓ *Creator*,
- ✓ *Date (Publication)*,
- ✓ *Type*,
- ✓ *Identifier*,
- ✓ *Relation (Project Id)* (il numero di *Grant Agreement*),
- ✓ *accessRights* (il tipo d'accesso),
- ✓ *embargoEndDate* (la data del periodo dell'embargo che deve essere rispettata dopo la pubblicazione effettiva dei lavori di ricerca sulle Riviste Scientifiche "tradizionali").

"Linee Guida per la creazione e la gestione di metadati nei Repository Istituzionali":

<<http://www.openaire.eu/en/component/content/article/9-news-events/427-openaire-releases-version-20-of-the-openaire-guidelines>>

Queste Linee Guida sono redatte in conformità alle Linee Guida DRIVER e OpenAIRE, entro la Commissione CRUI (gruppo *Open Access*). Hanno l'obiettivo di suggerire come descrivere i lavori di ricerca entro gli Archivi Istituzionali attraverso uno *schema di metadati DC minimo*, per provvedere a una maggiore interoperabilità tra i contenuti di diversi Archivi Istituzionali.

"L'interoperabilità è la capacità, posseduta dai sistemi multipli, di scambiarsi dati con una perdita minima di contenuto e funzionalità, nonostante i sistemi multipli siano caratterizzati da *Hardware*, piattaforme di *Software*, strutture di dati e interfacce differenti" [V. sito "Linee Guida per la creazione e la gestione di metadati nei Repository Istituzionali" p. 11].

Lo schema di metadati proposto dalle "Linee" fornisce i metadati di base (descrittivi e

strutturali) e quelli legati al contesto di appartenenza di un documento (es. informazioni amministrativo-gestionali relative ad afferenza dipartimentale, Settori Scientifico-Disciplinari: SSD) in grado di garantire la massima visibilità e la disseminazione alla produzione scientifica di un Ateneo, anche nell'ambito dell'Anagrafe Nazionale delle Ricerche (Sito Docente M.I.U.R).

Viene prevista la mappatura dei metadati con il modello di metadati per i lavori accademici SWAP (*Scholarly Works Application Profile*).

Viene preso in considerazione il modello di dati di CERIF (*Common European Research Information Format*) <<http://cordis.europa.eu/cerif/src/about.htm>>, uno Standard per favorire lo scambio delle informazioni in diversi formati tra i vari CRIS (*Current Research Information System*).

“Il fatto che i *Software* maggiormente utilizzati per la gestione di tali *Repository* utilizzino all'interno della Banca Dati dei formati specifici, collegati alla struttura stessa del Database, permetterà senza grosse difficoltà l'esposizione di metadati di formati diversi, a seconda delle diverse necessità, a partire dai medesimi dati descrittivi” [ivi, p. 31].

Metadati amministrativi e gestionali

Questi metadati [V. sito “MAG Metadati Amministrativi e Gestionali“], più che quelli descrittivi, assumono un'importanza preponderante ai fini della gestione, del mantenimento e dell'accessibilità a lungo termine degli oggetti digitali, in quanto documentano le modalità di *generazione, immissione, archiviazione e manutenzione* dei medesimi e forniscono specifiche formali per le fasi di raccolta e archiviazione dei metadati [V. sito “*Reference Model for an Open Archival Information System: OAIS*“].

Lo schema di metadati METS (*Metadata Encoding and Transmission Standard*) [V. sito]. “è uno standard per codificare metadati *descrittivi, amministrativi e strutturali* relativi a oggetti all'interno di una Biblioteca Digitale, esattamente come MAG, ma che, a differenza di MAG, si propone più come “*schema contenitore*” che come risposta o punto di riferimento per la *registrazione di metadati*. METS può integrare al suo interno diversi schemi di codifica non predeterminati, mentre da ben poche indicazioni su come e cosa si debba codificare” [V. sito “MAG“].

Ogni collezione deve essere descritta da un *file METS specifico* che contiene al suo interno i *riferimenti* utili ad *individuare gli oggetti* di appartenenza e le *informazioni di sintesi* per ogni risorsa.

Secondo le regole definite nel profilo METS, l'oggetto informativo digitale ospitato da una piattaforma deve:

- ✓ avere il proprio identificativo univoco (secondo la buona pratica, un *PID/Persistent Identifier* che garantisce all'oggetto l'unicità e la persistenza a livello di *Repository*). PID deve permettere la generazione di oggetti di livello gerarchico inferiore. In questo modo, gli oggetti “figli”, oltre ad avere una relazione che li lega all'oggetto padre, potranno avere un identificativo ricavato gerarchicamente dall'identificativo del padre. Gli oggetti “padre” ospitati entro la piattaforma *Software Open Source Fedora* sono in *relazione* con gli oggetti “figli” sulla base di *relazioni RDF* che descrivono anche la semantica della relazione stessa (modellazione atomistica);
- ✓ essere individuato nel sistema da un *file METS* avente il medesimo identificativo e che contiene tutti i metadati utili e le informazioni necessarie per la corretta esecuzione dell'*ingestion* (caricamento) dell'oggetto, la collezione di oggetti digitali e i loro metadati;
- ✓ appartenere necessariamente ad una collezione (solo se la risorsa deve essere indicizzata e visionabile da un'applicazione *client* mentre in caso di sola conservazione l'appartenenza ad una collezione non è vincolante);

- ✓ essere riconosciuto dal sistema come appartenente ad una specifica tipologia tra quelle definite dall'Archivio e come tale trattato per l'attivazione di specifici servizi.

Standard di metadati e protocolli raccomandati per gli Archivi digitali:

<<http://62.204.194.27/necobelac/generalita.jsp?id=id8752>> (Figura 30)

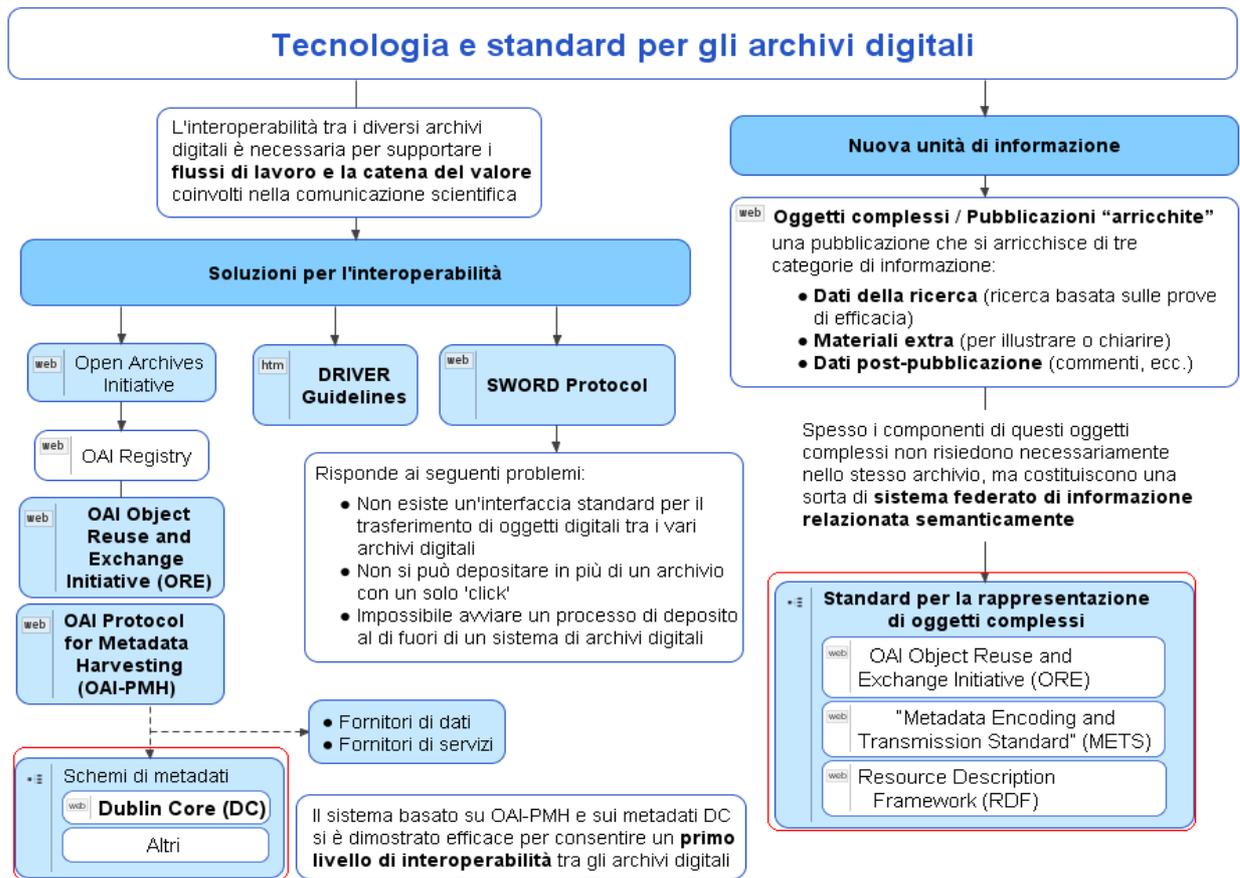


Figura 30. Tecnologia e standard per gli Archivi digitali

Strettamente connessi tra di loro sono i temi della *leggibilità* del *Repository* e della capacità dell'utente di *utilizzarlo*. Da una parte è necessario informare, *formare* e *sensibilizzare l'utente*, che nel caso degli Archivi Aperti può essere anche un *produttore* del *record* [scheda di metadati descrittivi della risorsa depositata]

Un utente conscio del funzionamento di questi particolari sistemi avrà meno difficoltà a COMPILARE i METADATI [Guerrini, 2010a, p. 55].

tramite “una maschera di immissione semplice ed intuitiva, magari con una procedura guidata passo per passo, a portata di click” [V. sito: “Linee guida Archivi Istituzionali”, p. 18]

necessari al momento del deposito di un contributo e alla creazione del suo *record*, poiché egli stesso avrà utilizzato l'Archivio anche come strumento di *DISCOVERY* dell'informazione [Guerrini, 2010a, p. 55].

4.3. Controllo di autorità dei valori di metadati

“Name Authority control is only one piece of the metadata-quality puzzle, but a piece worth examining” [Salo, 2009, p.1].

Analogamente alle lingue naturali, anche i linguaggi artificiali con i quali viene realizzata la mediazione tra una raccolta di documenti ed i suoi potenziali utenti sono costituiti [...] da quattro componenti necessarie:

- ✓ il VOCABOLARIO, cioè l'insieme delle espressioni elementari utilizzate per denominare le singole entità, i singoli attributi e le singole relazioni;
- ✓ la SEMANTICA, che concerne il significato per il quale una determinata espressione è inclusa nel vocabolario ed impiegata nel linguaggio, e quindi le *relazioni a priori* che derivano da quel significato e lo definiscono;
- ✓ la SINTASSI, che concerne la formazione, mediante ordinamento di singoli elementi tratti dal vocabolario, di espressioni più complesse;
- ✓ la PRAGMATICA, che attiene alle condizioni e modalità di applicazione del linguaggio [...]

La semantica referenziale concerne l'unicità di significato (univocità) di ciascuna delle espressioni elementari comprese nel vocabolario: è compito del Controllo d'Autorità garantire questa unicità, corredando ogni espressione di tutti gli elementi formali strettamente indispensabili e sufficienti ad impedire qualsiasi equivoco, di uso e di interpretazione” [Tartaglia, 2003, p.4].

Il *linguaggio di indicizzazione controllato* che l'indicizzatore usa per interpretare e rappresentare tematiche e concetti:

- ✓ e' un corpus di conoscenze rappresentate da un linguaggio specifico di un particolare argomento o area disciplinare [definiti da] collezioni di parole e di frasi, chiamate Terminologie, strutturate in modo da mostrare un ricco insieme di relazioni semantiche tra termini e concetti;
- ✓ deve servire allo scopo di ridurre le distanze tra le domande degli uomini e le risposte dei sistemi, facilitando *l'analisi dei documenti* e il loro conseguente *recupero*.

Il processo che conduce al Vocabolario controllato permette sia all'indicizzatore che a chi cerca di accedere ai medesimi concetti attraverso *termini autorizzati*, noto come 'descrittori'.

Il *Vocabulary Control* può essere supportato da una diversità di sistemi di organizzazione della conoscenza (*Knowledge Organization Systems, KOSs*)⁷ come *Thesauri, Terminologie, Classificazioni, Schemi a intestazione di soggetto, Registri dell'autorità dei nomi* ampiamente riconosciuti e condivisi, raccordando così anche diversi sistemi nelle operazioni di descrizione e di percorsi di ricerca e *browsing* delle risorse informative.

Nell'ottica di condivisione dei dati si deve cercare di *promuovere e condividere* le LISTE di AUTORITÀ con i termini controllati e altri KOS tra più sistemi e utenti nel contesto della Rete, caratterizzato dall'apertura e dalla necessità di convivenza con molti sistemi diversi, realizzando la tanto agognata *interoperabilità* sia tra diversi sistemi informativi digitali che tra i contenuti che essi ospitano [Gnolli, 2010 , presentazione].

Le procedure che hanno largamente beneficiato dell'informatizzazione e del Web sono le seguenti:

- ✓ Liste d'autorità (*AUTHORITY FILE*): *compilazione di termini* o di intestazioni autorizzate utilizzate da una singola organizzazione o da un consorzio per la catalogazione, l'indicizzazione o la documentazione [...]
- ✓ Controllo d'autorità (*AUTHORITY CONTROL*): un sistema di procedure che conserva informazioni consistenti nel *RECORD di un database*. Le procedure includono la *registrazione dei termini* e la loro *validazione* utilizzando *Liste di autorità*, con l'obiettivo di consentire a chi cerca di collegare tra loro termini simili [...] Questo lavoro può includere: verifica dei termini o dei nomi proposti in fonti autorevoli come *dizionari, monografie o fonti storiche*; ricerca di sinonimi come *varianti fonetiche*;

⁷ Dei KOS e degli approcci semantici all'informazione si occupa un'associazione internazionale *International Society for Knowledge Organization (ISKO)*. ISKO Italia è un capitolo nazionale dell'associazione: <<http://www.iskoi.org/>>

definizione di relazioni tra termini o nomi; creazione di record di autorità [*AUTHORITY RECORD*] da poter essere inseriti nella base di dati [...]

✓ L'*intestazione* è costituita da una *registrazione*, non solo da un nome, chiamata *AUTHORITY RECORD*, comprensiva della *FORMA STANDARD*, delle *FORME VARIANTI* e delle forme correlate in altre lingue [...] Ciascun *AUTHORITY RECORD* è abbinato a uno o a più record bibliografici“ [schede descrittivi]: in tal modo il Catalogo digitale consente di accedere alle registrazioni delle opere di un Autore da qualsiasi forma correlata e da qualsiasi ordine di combinazione”

[Innocenti, pp.107-108].

L'uso di una forma controllata permette

all'utente di risparmiare tempo e fatica grazie alla esplicitazione dei RINVII alle FORME CONTROLLATE delle intestazioni e al raggruppamento delle opere sotto una sola forma [...]

a partire dall'esistenza di *AUTHORITY FILE* internazionali, reali o virtuali, in cui siano registrate le varianti del NOME di ciascuna entità e le forme parallele nelle diverse lingue e nei diversi alfabeti, si possono sviluppare delle *interfacce adatte* alle esigenze e alle scelte di ciascun utente

[Innocenti, 2004, pp.106-107, 109].

Un sistema di *AUTHORITY CONTROL*, i cui *record* siano in grado di accogliere tutti i dati occorrenti a far sì che essi possano fungere da punto di accesso agli altri sistemi collegati, potrà consentire la NAVIGAZIONE TRASVERSALE. Per realizzare questo sistema, occorrerà che l'*AUTHORITY FILE* venga alimentato dall'attività di catalogazione delle istituzioni coinvolte, prelevando, per così dire, da ciascuna determinati segmenti di informazione, che poi troveranno nella notizia del sistema di *AUTHORITY CONTROL* una ricomposizione funzionale [Weston, Galeffi, 2003, p.8]

Nei sistemi informativi digitali

L'*autorevolezza della fonte* diventa l'elemento discriminante, la selezione avviene a monte, preferendo *scegliere la risorsa* sulla base dell'*autorevolezza* di chi l'ha creata, anziché a valle, scegliendo acriticamente sulla base del *ranking* [dei risultati] che appaiono sulla pagina. La qualità della fonte, la certezza della provenienza diventano, dunque, elementi determinanti nel percorso esplorativo del ricercatore [Guerrini, Possemato, 2013, pp.9-10].

Gli IR nascono senza perseguire l'obiettivo di gestire forme di *controllo di autorità* sui metadati dei contributi depositati, ma è evidente che, per esempio, *l'assenza del Controllo* della forma del Nome dell'Autore rende *difficoltoso il reperimento* degli *item* prodotti da un singolo Autore e *ostacola l'aggregazione delle opere*. È sufficiente compiere una ricerca in Internet per avere un'idea delle difficoltà che possono incontrare gli utenti. Le cause sono diverse [...] auto-archiviazione e auto-inserimento dei metadati da parte degli autori; procedure di inserimento dei dati in modalità automatizzata (*batch*); procedure di inserimento di dati archiviati (*backfiles*) per i quali *non sono state controllate* le eventuali *forme varianti* del nome dell'autore [Guerrini, 2010 a, p.46].

Il processo dell'*AUTHORITY CONTROL* deve essere posto al servizio del *Repository*:

- ✓ per assicurare l'omogeneità formale di ciascuna voce scelta come accesso (Autore, Titolo, Ente, Soggetto);
- ✓ per evitare la proliferazione di errori (ambiguità a cascata);
- ✓ per risolvere il problema degli omonimi e dei sinonimi, con ricaduta apprezzabile sull'attendibilità dei dati citazionali;
- ✓ per facilitare la scoperta, l'*identificazione*, la *selezione* e l'*uso delle risorse digitali* da parte degli utenti finali;
- ✓ per *agevolare i processi amministrativi*, così come l'identificazione della *provenienza*, la *gestione dei diritti*, e la *conservazione* delle risorse.

In seguito saranno esposte informazioni, elencate fonti, e forniti rimandi di approfondimento relativamente al tema del presente paragrafo.

”Dichiarazione di Principi Internazionali di Catalogazione (IFLA)”:

<http://www.ifla.org/files/assets/cataloguing/icp/icp_2009-it.pdf>

- La Dichiarazione si fonda sulle grandi tradizioni catalografiche del mondo e sul modello

concettuale di *Functional Requirements for Bibliographic Records* (FRBR) dell'IFLA [...]

□ È auspicabile che la Dichiarazione possa incrementare la condivisione internazionale di dati bibliografici e di Autorità e guidare gli estensori di regole di catalogazione nel loro impegno per l'elaborazione di un Codice di catalogazione internazionale" [...]

□ Possono essere rappresentate con dati bibliografici e di Autorità le seguenti entità:
Opera – Espressione – Manifestazione - Esemplare - Persona – Famiglia – Ente – Concetto – Oggetto – Evento – Luogo [...]

□ I Punti di ACCESSO CONTROLLATI forniscono la coerenza necessaria per raggruppare le registrazioni bibliografiche per insiemi di risorse.

“Linee Guida per gli Archivi Istituzionali” (CRUI, Italia):

<<http://www.cru.it/HomePage.aspx?ref=1781>>

□ La ricerca può avvenire tramite LINGUAGGI CONTROLLATI (secondo le notazioni di vocabolario/i controllato/i e altri sistema/i di organizzazione della conoscenza, e/o Banche di Dati di autorità scelto/i) e tramite parole o espressioni libere.

□ L'integrazione dell'Archivio tramite interfacce Standard e *Web Services* con Banche Dati Certificate d'Ateneo (es. Anagrafe della Ricerca; Sistema di autenticazione/autorizzazione centralizzata via *LDAP*, *CAS*, *Shibboleth dell'Ateneo*; Sito Docente del Ministero) e con Server su cui vengono caricati i vocabolari controllati e altri KOS è un passo importante per:

- ✓ la validazione della qualità dei Metadati al suo interno;
- ✓ pulizia degli indici di ricerca;
- ✓ potenziare l'interoperabilità dell'Archivio con altre applicazioni;
- ✓ incentivare il radicamento dell'Archivio nelle abitudini di studio e ricerca degli utenti autori.

Virtual International Authority File (VIAF): <<http://viaf.org/>>

È un File di Authority internazionale mantenuto da OCLC.

VIAF è disponibile in modalità *Linked Open Data* (Figura 31) ha una grande importanza per la gestione e la rappresentazione dei Nomi controllati di autori (di libri) in diverse lingue. *Linked Open Data* aumenta l'utilità delle informazioni sul web incrociando riferimenti a persone, luoghi, cose, più consistenti e collegabili attraverso domini.

The screenshot shows the VIAF website interface. At the top, it says 'VIAF Virtual International Authority File' and '19 headings found for ferdinand de saussure'. Below this is a table with three columns: 'Heading', 'Type', and 'Sample Title'. The first row shows a heading for 'Saussure, Ferdinand de 1857-1913' with various language flags, a 'Personal' type, and a sample title 'Cours de linguistique générale.' in French. The second row shows the same heading in German: 'Saussure, Ferdinand de 1857-1913', with German flags, the same 'Personal' type, and the same sample title in German: 'Cours de linguistique générale.'

Heading	Type	Sample Title
1 Saussure, Ferdinand de 1857-1913	Personal	Cours de linguistique générale.
Saussure, Ferdinand de 1857-1913	Personal	Cours de linguistique générale.

Figura 31. Pagina online del *Virtual International Authority File* (VIAF)

□ VIAF è concepito per un accesso paritario multilinguistico, laddove a confronto non sono tanto lingue diverse, ma convenzioni nomenclatorie particolari, consolidate e, nei differenti abiti, ugualmente pertinenti.

□ VIAF “permette di collegare tra loro le registrazioni di autorità relative a una medesima entità usate nei vari paesi e dunque di realizzare la funzione di raggruppamento. Esso potrebbe rivelarsi molto utile anche per i *Repositories*” [Guerrini, 2010 b, pp. 6, 7].

Getty Union of Artist Names (ULAN):

<<http://www.getty.edu/research/tools/vocabularies/ulan/index.html>>

È un File di Authority internazionale con più di 225.000 Nomi (incluso pseudonimi, varianti e varianti linguistiche) con le informazioni biografiche su artisti e architetti. È incluso nel VIAF.

“Author Service” dell’aggregatore degli Archivi digitali in Economia RePEc:

<<http://authors.repec.org/about>>

Questo servizio fornisce le voci di autorità per i Nomi di ricercatori in scienze economiche registrati nel RePec.

Servizi per l’identificazione e il controllo dei Nomi propri di Autori

CrossRef: article-linking consortium

<http://www.crossref.org/CrossTech/2007/02/crossref_author_id_meeting.html>Questo servizio attribuisce gli identificatori univoci (“Author IDs”) agli autori che pubblicano presso alcuni editori.

Research Crossroads: <<http://www.researchcrossroads.org/>>

Questo servizio attribuisce ad ogni autore/ricercatore registrato nel sistema un identificatore univoco “ResearchCrossroads ID”.

ISNI (International Standard Name Identifier): <<http://www.isni.org/>>

Questo servizio identifica creatori di opere con un unico numero a 16 cifre.

ORCID (Open Researcher and Contributor ID): <<http://www.orcid.org>>

Questo servizio ha lo scopo di risolvere l’ambiguità sistemica dei Nomi e - mediante l’assegnazione di identificatori unici abbinabili ai risultati della ricerca di un ricercatore - di migliorare il processo di scoperta dell’informazione scientifica e di migliorare l’efficienza del finanziamento e la collaborazione.

International Registry of Authors Links to Identify Scientists (IRALIS): <<http://www.iralis.org/>>

Il progetto IraLIS è stato avviato per risolvere il problema di disambiguazione dei Nomi in paesi di lingua spagnola. Analoghi problemi di identificazione dell’Autore appaiono in tutte le lingue a causa delle forme varianti di voci dei Nomi.

Qualsiasi persona che opera in un campo della ricerca scientifica può registrarsi in IraLIS creando un Record per il proprio Nome, a prescindere se ha già pubblicato qualche lavoro scientifico o prevede di farlo in futuro.

IraLis è basato su un’ontologia (disponibile sul web come *Linked Open Data*) comprensibile dal computer atta a descrivere persone, con le loro attività e le relazioni definite dal vocabolario FOAF (*Friend of a Friend - Amico di un amico*). I record IRALIS vengono indicizzati da *Google*, *Google Scholar*, *OAIster*, *Recolecta*.

Elenco di alcuni importanti Sistemi per l’identificazione e il controllo dei Nomi propri di Autori:

<<https://repinf.pbworks.com/w/page/13779410/Author%20identification>>

Disambiguazione dei Nomi degli Autori e dei Titoli delle Riviste

offerta dal Software Eprints: <<http://www.eprints.org/software/v3>>

Le ultime versioni di *EPrints* incorporano i moduli che permettono:

compiere sondaggi nella *Lista degli Autori* caricati nel sistema (*Repository*) da Database dell’Anagrafica del personale dell’Istituzione;

suggerire un Nome dell’Autore (Autore, Curatore, Autore di prefazioni) al depositante del materiale nel sistema per il completamento automatico di qualità dei valori del metadato “Autore”;

la gestione di un *File di Authority* completo per i Titoli delle Riviste scientifiche,

derivato da *SHERPA/RoMEO* (la banca dati delle politiche sul Copyright e l'archiviazione dei diversi Editori) [V. sito].

Suite di servizi applicativi per il governo (*governance*) degli Atenei e degli Enti di ricerca

SURplus (il Sistema Universitario Ricerca plus): <<http://surplus-unibic.cilea.it/>>;

U-GOV: <<http://www.cineca.it/it/content/il-sistema-u-gov>>

□ Una volta sono integrati con gli Archivi Istituzionali, questi servizi di *governance* permettono alle piattaforme degli Archivi di collegarsi dinamicamente con vari sistemi gestionali dei rispettivi Atenei, consentendo in fase di inserimento di metadati di reperire i Nomi (di Enti, di Persone) dalle Anagrafi ufficiali, con indubbi vantaggi per la qualità di ricerca dell'informazione. L'utilizzo delle Anagrafiche comuni a tutto lo strato applicativo (struttura condivisa delle informazioni) consente di condividere i dati attraverso diversi processi, evitando la duplicazione di informazioni.

□ I metadati sviluppati sui moduli SURplus *Open Archive (OA)* e UGOV-Ricerca, tramite il protocollo dell'interoperabilità OAI-PMH hanno uscite per il "Sito Docenti" del MiUR. Ciò è possibile grazie all'associazione di determinati metadati (es. "Soggetto", "Disciplina", "Classificazione") con valori della Classificazione MiUR - Settori Scientifico-Disciplinari (SSD).

Schemi di Classificazione

le Classificazioni di Soggetto possono dar vita a strumenti per *l'organizzazione della conoscenza* per collezioni lessicali estratte da *database di metadati o documenti a testo completo*, per *Terminologie, Glossari, Dizionari o Enciclopedie, Rassegne*, fin anche a biblioteche distribuite di documenti digitali nativi o digitalizzati da supporto analogico. L'insieme delle descrizioni di uno Schema di Classificazione costituisce già di per sé una importante risorsa Terminologica [De Robbio, 2007, p.189]

□ **Classificazione MiUR - Settori Scientifico-Disciplinari (SSD):**

<<http://cercauniversita.cineca.it/php5/settori/index.php>>

È una Classificazione che raggruppa le discipline definite dal MIUR in 28 classi (Figura 32). Ad ogni Settore Disciplinare appartengono tutte le materie riconducibili alla medesima *declaratoria*, cioè quella sezione del Decreto che descrive sinteticamente i contenuti di ogni singolo Settore Scientifico.

Gruppi				MIUR Settori scientifico-disciplinari			
AGR	BIO	CHIM	FIS	Are 01 Scienze matematiche e informatiche 02 Scienze fisiche 03 Scienze chimiche 04 Scienze della Terra 05 Scienze biologiche 06 Scienze mediche 07 Scienze agrarie e veterinarie 08 Ingegneria civile ed Architettura 09 Ingegneria industriale e dell'informazione 10 Scienze dell'antichità, filologico-letterarie e storico-artistiche 11 Scienze storiche, filosofiche, pedagogiche e psicologiche 12 Scienze giuridiche 13 Scienze economiche e statistiche 14 Scienze politiche e sociali			
GEO	ICAR	INF	ING-IND				
ING-INF	IUS	L-ANT	L-ART				
L-FIL-LET	L-LIN	L-OR	M-DEA				
M-EDF	M-FIL	M-GGR	M-PED				
M-PSI	M-STO	MAT	MED				
SECS-P	SECS-S	SPS	VET				

Figura 32. Pagina online della Classificazione del MiUR - SSD (Settori Scientifico-

Disciplinari)

Una gran parte degli Archivi Istituzionali si serve di questi descrittori, utili più alle statistiche (per il MiUR) e ai processi di valutazione, a livello d'Ateneo e su scala nazionale, che a favorire il reperimento delle risorse documentarie, tanto più che spesso sono riferiti al Docente piuttosto che alle risorse che dovrebbero descrivere. Sono significativi, inoltre solo per una parte della comunità degli Autori italiani, gli universitari, e sono assolutamente incomprensibili per qualunque straniero. In un ambiente globale ciò è ovviamente insostenibile [...]

Di converso, è certo che per la comunità accademica italiana si tratti di dati fondamentali, soprattutto all'interno di analisi su base nazionale. Sono, forse, l'unico dato veramente condiviso, tanto è vero che vengono usati per la *ricerca per argomento* dal *Service Provider* di PLEIADI [Guerrini, 2010 b, p. 10]

□ **Classificazione Decimale Dewey (DDC):**

<<http://www.oclc.org/dewey/webservices.en.html>>

- ✓ È un sistema (pubblicato online anche in modalità *Linked Open Data*: <<http://dewey.info/>> per la divisione del sapere in *dieci Classi* (attraverso le cifre decimali secondo i principi della *disciplina* e della *gerarchia*) maggiormente usato a livello internazionale (Figura 33).
- ✓ Le Classi si dividono a loro volta in *dieci divisioni* anch'esse numerate da 0 a 9, e ogni divisione in *dieci sezioni* ancora numerate da 0 a 9.

List of Dewey Decimal classes

- 000 – Computer science, Library and Information science & general work
- 100 – Philosophy and psychology
- 200 – Religion
- 300 – Social sciences
- 400 – Language
- 500 – Science
- 600 – Technology
- 700 – Arts
- 800 – Literature
- 900 – History, geography & biography

Figura 33. Pagina online della Classificazione Decimale Dewey

□ **Universal Decimal Classification (UDC):**

<<http://www.udcc.org/index.php/site/page?view=factsheet>>

- ✓ È uno schema di Classificazione multilingue ispirato da DDC, ma è molto più potente (Figura 34).

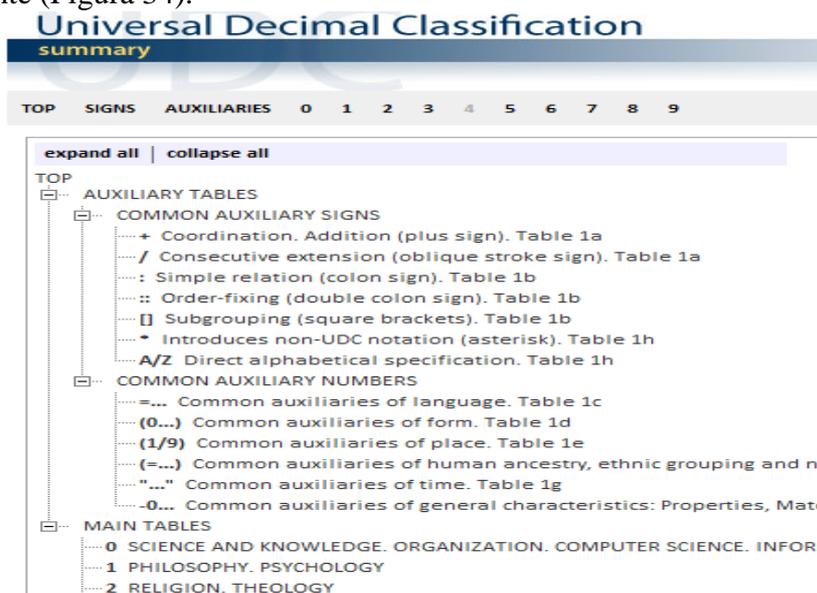


Figura 34. Pagina online della Classificazione Universale Decimale

- ✓ UDC si serve di segni ausiliari per indicare vari aspetti particolari di un Soggetto e delle Relazioni fra soggetti. In tal modo contiene un elemento significativo *faccettato* o analitico-sintetico che la rende adatta soprattutto nelle Biblioteche Specialistiche.
- ✓ UDC consente di classificare documenti di ogni genere: innanzi tutto documenti della letteratura, cioè *documenti scritti*, ma anche su altri media come *registrazioni su pellicola, video, registrazioni sonore, illustrazioni, mappe e artefatti* (realia) come pezzi da museo.
- ✓ La CDU per gli identificatori dei soggetti si serve della numerali arabi e di notazioni del sistema decimale.
- ✓ Il *dataset "UDC Summary"* rappresenta una selezione di circa 2000 classi estratti dallo schema UDC e pubblicati come *Linked Open Data*:
<<http://www.udcc.org/udcsummary/php/index.php>>

□ **Classificazione della *Library of Congress* (LCC):**

<<http://www.loc.gov/catdir/cpsol/lcc.html>>

- ✓ È un sistema di 21 classi di organizzazione della conoscenza che assegna ad ogni Classe una lettera dell'alfabeto in cui potrebbe trovare la collocazione un Soggetto (tema) che descrive una risorsa (Figura 35).
- ✓ È utilizzato da gran parte delle Biblioteche degli istituti di ricerca e universitari degli Stati Uniti e da molte biblioteche simili in altri paesi, anche se numerose biblioteche pubbliche di dimensioni medio-grandi continuano a servirsi della Classificazione decimale Dewey (DDC).

Library of Congress Classification

- ✦ [A -- GENERAL WORKS - WP version - Word version](#)
- ✦ [B -- PHILOSOPHY, PSYCHOLOGY, RELIGION - WP version - Word version](#)
- ✦ [C -- AUXILIARY SCIENCES OF HISTORY - WP version - Word version](#)
- ✦ [D -- WORLD HISTORY AND HISTORY OF EUROPE, ASIA, AFRICA, AUSTRALIA, NEW ZEALAND, ETC. - WP version - Word version](#)

Figura 35. Pagina online della Classificazione di *Library of Congress*

□ **COLON Classification:** <<http://www.aib.it/aib/contr/lavazza1.htm>>

- ✓ È uno schema molto diffuso in India per Biblioteche pubbliche e di ricerca. Molti sistemi di classificazione hanno tentato di recepirne gli insegnamenti.
- ✓ Racchiude l'intero universo di conoscenze in 40 discipline o campi di studio, corrispondenti alle Classi principali dello schema.
- ✓ La notazione è mista di *numeri, lettere latine o greche e speciali operatori di calcolo* (ad esempio z =generalia, 2 =library science, AZ =mathematics, NX =literature and languages, Σ =social science).
- ✓ In ognuna delle Classi della colon classification, hanno notazione propria solo i soggetti isolati, semplici, uniti tra loro in gruppi omogenei.
- ✓ I soggetti composti non vengono enumerati.
- ✓ I nuclei di *concetti affini*, individuati all'interno di ciascuna classe, sono denominati *faccette*; all'interno di ogni *faccetta*, tutti i *concetti isolati* (foci) avranno *in comune una particolare caratteristica*.
- ✓ Per classificare un soggetto composto occorre, individuata la *Classe principale*, analizzarlo in *parti elementari*, ricercare attraverso gli indici delle *tavole le*

notazioni dei singoli elementi e, infine, ricombinarli, in base alle reciproche relazioni, per riformare il *Soggetto Composto* espresso in linguaggio notazionale *Colon Classification*.

- ✓ Per evitare una "fuga all'infinito" del numero delle possibili relazioni, *Colon Classification* è ridotta a cinque categorie fondamentali di concetti isolati, applicabili a qualsiasi area del sapere. La formula di combinazione di tali categorie è riassunta nell'abbreviazione PMEST (*PERSONALITY*: personalità o entità o cosa; *MATTER*: materia o materiale o proprietà; *ENERGY*: energia o processo, operazione, azione; *SPACE*: spazio); *TIME*: tempo).

□ **JITA Classification of Library and Information Science (LIS):**

<<http://datahub.io/dataset/jita>>

- ✓ Attualmente è disponibile in Inglese e in altre 14 lingue (*Arabic, Catalan, Czech, Dutch, English, French, Greek Modern (1453-), Italian, Polish, Portuguese, Romanian, Russian, Spanish, Swedish, Turkish*);
- ✓ È anche accessibile come *Linked Open Data*;
- ✓ Rappresenta le voci di "Soggetto" nell'Archivio internazionale aperto per i documenti scientifici in biblioteconomia e scienza dell'informazione "E-LIS" [V.sito]

La ricerca per Classe permette di *rintracciare opere* che presentano *proprietà o caratteristiche comuni* secondo un ordine logico sistematico stabilito in precedenza da schemi di Classificazione.

□ **BLISS Bibliographic Classification (a Classification for the 21st century):** <<http://www.aib.it/aib/contr/broughton1.htm>> [Broughton, 2004].

È una Classificazione bibliografica interamente basata su uno *schema a faccette*.

Library Congress Subject Headings (LCSH): <<http://id.loc.gov/search/>>

È un sistema (disponibile online anche in modalità *Linked Open Data*) dell'intestazione di Soggetto attraverso le voci gerarchiche predefinite a monte.

□ Mappatura dei valori di LCSH con i numeri della DDC:

<<http://www.oclc.org/dewey/updates/numbers.en.html>>

□ **FAST (Faceted Application of Subject Terminology):** <<http://fast.oclc.org>>

FAST nell'OCLC Terminologies Service: <<http://tspilot.oclc.org/resources>>

□ FAST rappresenta un vocabolario controllato molto ricco di termini del sistema LCSH con l'obiettivo di renderli facilmente comprensibili, applicati e utilizzati in diversi sistemi di gestione dell'informazione, sia allo scopo del *Controllo d'autorità* dei valori di metadati che qualificano l'intestazione di *Soggetto* (tramite 1.600.000 Record di autorità), che per contribuire ad una *migliore interoperabilità semantica* tra i medesimi valori.

□ Tutte le intestazioni di Soggetto sono opportunamente raccordate da *Record di Autorità* eliminando la necessità di sintetizzare le intestazioni insieme alle regole sintattiche.

□ I singoli termini del vocabolario FAST sono distribuiti in 8 categorie distinte (*faccette*): *Nomi personali, Nomi aziendali, Nomi geografici, Eventi, Titoli, Periodi di tempo, Tematiche, Forma /Genere*.

□ FAST è disponibile online anche come *Linked Open Data*: <<http://id.worldcat.org/fast/>> (attraverso la *Open Data Commons Attribution Licence ODC-By v1.0*): <<http://www.oclc.org/research/activities/fast/download.html>>

Thesaurus del NUOVO SOGGETTARIO della Biblioteca Nazionale Centrale di Firenze (BNCF): <<http://thes.bncf.firenze.sbn.it/>>

□ È un sistema dell'intestazione di Soggetto che propone un linguaggio pre-coordinato, sia *analitico* che *sintetico*, nel rispetto delle norme internazionali in materia di controllo e struttura del vocabolario basato sull'analisi categorica di relazioni semantiche.

□ È creato in vista di un *Thesaurus universale (multidisciplinare)* che, a regime, arriverà a contenere circa 70.000 voci in continuo accrescimento, utilizzabili nell'indicizzazione di materiali di varia natura. Prevede una netta distinzione tra relazioni semantiche e quelle sintattiche, e basa il suo *ordine citazionale delle stringhe di Soggetto* sul modello di analisi per le relazioni logiche.

□ È disponibile online tramite i formalismi del Web Semantico per la pubblicazione dei sistemi di organizzazione della conoscenza (KOS) - *Simple Knowledge Organization System (SKOS)*.

□ Ha l'obiettivo di

valorizzare le potenzialità del linguaggio sia sul versante dell'indicizzazione *pre-coordinata* (costruzione di *stringhe di soggetto* con l'utilizzo di *terminologia controllata*), sia *post-coordinata* (impiego esclusivo di parole chiave coincidenti con la terminologia controllata) [...]

I vantaggi di utilizzare anche nel contesto degli IR un linguaggio del genere, diffuso a livello nazionale e condiviso da numerosi Centri Documentari, sarebbero molti: il riferimento a un *lessico affidabile* costruito secondo Standard Internazionali e fortemente legato alla garanzia bibliografica [...] ma soprattutto l'*abbandono* di *scelte soggettive* e fantasiose che, se basate sull'impiego di una *terminologia non controllata*, ostacolano qualsiasi forma di *cooperazione e condivisione* nel reperimento efficace ed esaustivo di risorse documentarie che trattano lo stesso argomento [Guerrini, 2010a, p.48]

Authority File attraverso OAI-feeds

□ "Nel rapporto tra *Data Provider*, *Service Provider* e *AUTHORITY CONTROL* diventano molto indispensabili gli IDENTIFICATORI UNIVOCI delle risorse digitali come URI e DOI, la cui missione è quella di non far duplicare le informazioni sui sistemi eterogenei in dialogo, e per collegare tra di loro documenti ospitati su sistemi differenti.

□ La presenza di identificatori, nella logica degli IDENTIFICATORI PERSISTENTI utilizzati per altre risorse disponibili in Rete - come, ad esempio, ISNI (*International Standard Names Identifier*), DOI, *handle* - permette l'aggregazione, a livello nazionale e internazionale, dei dati prodotti e aggiornati localmente.

□ Gli identificatori, o "uncici", rappresentano pertanto una *parte essenziale del metadato*, che, nel caso specifico degli Archivi Istituzionali, svolgono una *funzione strategica*, perché si occupano del *discovery* delle risorse e degli agganci nel mondo commerciale (motori di ricerca) e in quello libero *Open Access*.

□ Nell'ambito dell'OAI (*Open Archives Initiative*) gli IDENTIFICATORI UNIVOCI possono essere predisposti dai gestori di un determinato *AUTHORITY FILE* attraverso i cosiddetti *OAI-feeds*.

Nel caso di un aggiornamento dell'*AUTHORITY FILE* della fonte, le nuove informazioni possono essere individuate e catturate dagli utilizzatori dell'*OAI-Server* in maniera tempestiva e precisa, evitando ogni ambiguità.

OAI-feeds viene utilizzato, ad esempio, per l'aggiornamento del file di autorità VIAF (*Virtual International Authority File*).

Elenco di Vocabolari controllati online disponibili per il Web Semantico:

<<http://www.w3.org/2005/Incubator/ldd/XGR-ldd-vocabdataset-20111025/>>

La maggior parte di questi vocabolari sono disponibili in modalità *Linked Open Data*, con i

relativi casi d'uso

Functional requirements for Authority Data. A conceptual model. - L'edizione italiana
“**Requisiti funzionali per i DATI di AUTORITÀ. Un modello concettuale**”

<http://www.iccu.sbn.it/opencms/export/sites/iccu/documenti/ImpaginatoFRADper_internet.pdf>

A c. di ICCU (Istituto Centrale per il Catalogo Unico delle Biblioteche Italiane), 2010

Elenco dei Vocabollari controllati secondo la tipologia per domini specifici.

VEST Registry: <<http://aims.fao.org/vest-registry/browse-by-vocabularies%20>>

È un servizio online (Figura 36) da cui è possibile scegliere tra diversi tipi di vocabolari controllati e semantici secondo l'applicazione di 2 criteri: “*Domains*” e “*Types of Vocabulary*”:

VEST Registry. Vocabularies

Domains	Types of Vocabulary
Agriculture - General/All	-Encyclopedia
Animal Production and Health	-Glossary
Economics and Policy	-Lexical Database
Education and Extension	-Ontology
Engineering, Technology and Research	-Subject headings
Farming Practices and Systems	-Taxonomy
Fisheries and Aquaculture	-Terminology Database
Food safety and Human nutrition	-Thesaurus
Food Security	-Topic Tree

Figura 36. Pagina online del VEST Registry dei Vocabolari controllati per dominio e tipologia

La *normalizzazione* (l'attribuzione dei significati precisi e univoci) e l'*arricchimento* dei valori di metadati sul *versante semantico* attraverso le voci di Liste di autorità, vocabolari controllati e schemi semantici assicura garanzia formale della *qualità di metadati* dell'Archivio, in quanto contribuisce alla loro accuratezza, consistenza, sufficienza e affidabilità (“*accurate, consistent, sufficient, and thus reliable*”) [Greenberg et al., 2002, p.45]. I metadati di scarsa qualità possono significare che una risorsa rimarrà essenzialmente invisibile e inutilizzabile all'interno di un sistema in cui risiede.

Nella fase di inserimento dei metadati nell'Archivio è opportuno assistere chi compie l'auto-archiviazione tramite la predisposizione di *facilities* quali *Liste Dinamiche* per l'auto-completamento e la *cattura (importazione) dei metadati da fonti autorevoli* interne ed esterne, come, ad esempio:

- ✓ Anagrafica d'Ateneo per i dati dell'autore;
- ✓ Anagrafe e valutazione della ricerca, secondo il flusso che ogni Ateneo riterrà più opportuno;
- ✓ Banche Dati e repertori per i dati bibliografici;
- ✓ Vocabolari e Schemi semantici disponibili come *Linked Open Data* (es. VIAF, Classificazione Decimale Dewey e altri.)

È necessario che il personale addetto al supporto della gestione di metadati nell'Archivio pianifichi le operazioni di controllo dei valori di metadati implementati, di scoperta e correzione degli errori e, se è necessario, deve qualificare e validare ulteriormente i valori di metadati compilati dagli autori depositanti.

4. 4. Associazione di metadati con valori semantici

Ogni azione di valorizzazione di possibili soluzioni semantiche per la qualificazione dei valori di metadati in un Archivio di ricerca e riuso degli *strumenti di organizzazione della conoscenza* hanno

come presupposto *la corretta collocazione e accessibilità dei risultati*: requisiti preliminari a qualsiasi azione di disseminazione dell'informazione.

Il processo di associazione di metadati con valori semantici (*Semantic Authority Control* [Lucarelli, 2003]) tramite i sistemi semantico-lessicali come:

- ✓ Thesauri (collezioni di termini organizzati gerarchicamente e relazionati fra loro);
- ✓ Classificazioni (schemi della gestione delle conoscenze per Classi);
- ✓ Tassinomie (si può riferire sia alla Classificazione gerarchica di concetti, sia al principio stesso della Classificazione);
- ✓ Ontologie (entità concettuali) e loro relazioni per descrivere e rappresentare un'area della conoscenza. Le Ontologie web codificano la conoscenza di un certo ambito e quella trasversale ad altri ambiti, rendendole in tal modo condivise dalle applicazioni in più ambiti)

Avrà effetto diretto sulla creazione di accessi semantici ai contenuti dell'Archivio che devono essere presenti tra i filtri utili per la navigazione o per il *browsing* di Soggetti o Classi in ogni piattaforma di ricerca di qualità [Iacono, 2013].

La problematica dell'*Authority Control* coinvolge naturalmente, oltre agli *aspetti* descrittivi delle risorse documentarie, anche quelli *semantici*. Bisogna prevedere, oltre alle *keyword* (parole chiave) inserite dagli autori in fase di deposito, anche dei *termini qualificanti semanticamente* e forse *Crosswalk* che connettano tali *termini* a loro *equivalenti in altre lingue*, per poter assicurare una piena interoperabilità e un alto grado di reperibilità.

Si tratta di una problematica interessante, ancora tuttavia allo studio e sperimentata in alcune realtà, molto complessa. Non è infatti sempre possibile tracciare una precisa corrispondenza tra termini di lingue diverse. Tra i progetti che sperimentano questa possibilità vi è il *MACS Project, Enabling large-scale Multilingual Access to Subjects* [Guerrini, 2010 b, pp.9-10].

In seguito saranno esposte informazioni, elencate fonti, e forniti rimandi di approfondimento relativamente al tema del presente paragrafo.

“Linee Guida per gli Archivi Istituzionali” (CRUI, Italia):

<http://www.cru.it/HomePage.aspx?ref=1781>

□ “Anche se può rivelarsi dispendioso *l’inserimento e l’associazione di metadati per la Semantica* (Ontologie, Thesauri, Classificazioni etc.) [...] *informazioni arricchite* consentono di riutilizzare le informazioni dell’Archivio in una varietà di contesti (Anagrafe della ricerca, valutazione, repertorio delle competenze ecc.)” [ivi, p.3]

□ “più i *metadati sono specifici* più “costano”, ma ne guadagna l’Archivio in termini di *posizionamento*, in quanto i documenti depositati acquisiscono una migliore visibilità ed una maggiore disseminazione” [ivi, p.13]

CLASSIFICAZIONI. V. § 4.3 del presente Capitolo

- ✓ Decimale Dewey (DDC)
- ✓ *Universal Decimal Classification* (UDC)
- ✓ Classificazione della *Library of Congress* (LCC)
- ✓ *COLON Classification*
- ✓ *JITA Classification of Library and Information Science* (LIS)
- ✓ *BLISS Bibliographic Classification*
- ✓ Classificazione dei Settori Scientifico-Disciplinari (SSD) del MiUR

Tra i servizi che possono contribuire a elevare i livelli di qualità degli IR rientrano senza dubbio l’assegnazione di METADATI SEMANTICI ai contributi depositati e il potenziamento delle funzioni del *browsing* per SOGGETTO [...]

In Italia, attualmente, l’offerta di accessi semantici per questo tipo di materiale è molto diversificata. Alcuni IR consentono l’inserimento e pertanto la ricerca solo tramite la dichiaratoria dei SETTORI SCIENTIFICO-DISCIPLINARI [SSD] stabiliti dal MiUR, mentre altri IR stanno sperimentando l’impiego di SOGGETTARI elaborati e adottati da agenzie

catalografiche e pertanto redatti secondo un linguaggio documentario consolidato, oppure impiegano SCHEMI di CLASSIFICAZIONE di varia natura [Guerrini, 2010a, p. 49].

Library Congress Subject Headings (LCSH): <<http://id.loc.gov/search/>>

V. § 4.3 del presente Capitolo

L'uso delle LCSH è certamente positivo, perché consente di estrarre *metadati in inglese* che, oltre ad *assicurare garanzia formale*, rappresentano un *linguaggio di interscambio internazionale*. Tuttavia, la forte *consistenza numerica e multidisciplinare* delle *headings* previste dalle LCSH può costituire un elemento di *complessità per l'attribuzione di metadati semantici* agli IR

[Guerrini, 2010a, p. 48]

NUOVO SOGGETTARIO: <<http://thes.bncf.firenze.sbn.it/>>

V. § 4.3 del presente Capitolo

Dal momento che il *Nuovo Soggettario* prevede la possibilità di uno sviluppo multilingue e contiene già equivalenti inglesi scelti fra quelli attestati da LCSH, la creazione di un THESAURUS SPECIALIZZATO per gli IR, *derivato* dal Thesaurus del NUOVO SOGGETTARIO e con esso interoperabile, potrebbe rappresentare la migliore soluzione per uniformare il TRATTAMENTO SEMANTICO delle risorse documentarie istituzionali all'interno di Enti diversi, come pure per correlare fra loro IR e OPAC, strutture documentarie sicuramente raccordabili e che possono superare ogni forma di incomunicabilità [...]

In ambito IR, il confronto con la terminologia codificata del Thesaurus del Nuovo Soggettario potrebbe avvenire in modalità automatica o semi-automatica, per esempio tramite combinazione (*match*) fra parole:

- ✓ contenute nei documenti (*estratte sulla base della loro occorrenza*) e termini del Thesaurus, oppure fra parole
- ✓ scelte per gli *Abstract* e termini del Thesaurus.

Diversamente, l'attribuzione potrebbe avvenire da parte del depositante con l'*individuazione dei termini da abbinare* alla registrazione del *contributo depositato*, scelti fra quelli preferiti del *Nuovo Soggettario* e, presumibilmente, in coerenza con altri metadati già attribuiti a contributi che trattano dei medesimi temi

[Guerrini, 2010a, p. 48-49].

“Investigative Study of STANDARDS for Digital Repositories and Related Services”:

<<http://dare.uva.nl/document/93727/>>

Nei *Repository* della Germania l'organizzazione del contenuto è per lo più guidato dalla Classificazione per SOGGETTO, i valori di cui vengono allestiti attraverso la *coesistenza degli schemi di Classificazione multipli* (es. “*Decimal Dewey Classification*”: DDC; “*Deutsche National Bibliografie*”: DNB; “*Schlagwortnormdatei*”: SWD), mentre

Nei *Repository* dell'Inghilterra si usano per lo più le Classificazioni *Library of Congress Classification* e *SHERPA Oxford Classification*.

Da queste scelte si coglie un tentativo di *armonizzare la rappresentazione dei valori* del metadato SOGGETTO entro gli schemi di metadati adottati a livello nazionale.

Gli elementi del metadato SOGGETTO possono essere inoltre ricavati dalle Classificazioni dei settori disciplinari (es. “*Mathematics Subject Classification*”: MSC; “*Medical Subject Headings*”: MeSH) e da tante altre Classificazioni locali.

SKOS (Simple Knowledge Organization System): <<http://www.w3.org/2004/02/skos/>>; <<http://www.w3.org/TR/2009/PR-skos-reference-20090615/>>

V. APPENDICE A

SKOS rappresenta una metodologia standard per la produzione delle specifiche tecniche di alto livello per i sistemi di organizzazione della conoscenza KOS - affinché essi diventano armonizzati ed integrati negli altri standard e linguaggi propri del Web Semantico (ad es. nelle Ontologie OWL).

Usando SKOS, concetti possono essere:

- ✓ identificati con URIs;

- ✓ etichettati con le stringhe lessicali in una o più lingue;
- ✓ annotati con codici lessicali;
- ✓ documentati con vari tipi di note;
- ✓ linkati ad altri concetti e organizzati nelle gerarchie informali e associazioni di networks;
- ✓ aggregati negli schemi e raggruppati nelle collezioni ordinate;
- ✓ mappati ai concetti di altri schemi.

LOD (*Linked Open Data*):

<<http://www.w3.org/wiki/SweoIG/TaskForces/CommunityProjects/LinkingOpenData>><<http://dare.uva.nl/document/93727/>>

V. APPENDICE A

□ I dati collegati (*Linked Data*) rappresentano un'espressione pratica del Web Semantico.

Il termine *Linked Data* è usato per descrivere un metodo per *esporre, condividere e connettere* sul Web i *dataset* di diversi soggetti produttori tramite specifiche metodologie.

□ *Linked Data* e *Open Data* [De Robbio, 2012] si attengono all'ambito del Progetto *Linking Open Data*

(<<http://www.w3.org/wiki/SweoIG/TaskForces/CommunityProjects/LinkingOpenData>>) che riguarda la pubblicazione e il collegamento di diversi *Open Data* sul Web. Questi *Data* o *dataset* (oggetti informative digitali) devono essere:

- ✓ identificati da URIs (*Uniform Resource Identifiers*) appoggiati sul Protocollo Web HTTP;
- ✓ espressi e pubblicati sul Web tramite la sintassi ("*Subject-Value-Object*") del linguaggio del Web Semantico RDF (*Resource Description Framework*)

□ Grazie all'iniziativa del W3C *Linking Open Data* un numero sempre crescente di *Data Provider* di tutto il mondo ha cominciato di pubblicare e collegare i propri *dataset* sul Web secondo i principi *Linked (Open) Data* ((L(O)D)), contribuendo alla creazione di una comune base di conoscenza sul web.

Elenco di *Value Vocabularies* disponibili online come *Linked Open Data*, con i relativi casi d'uso: <<http://www.w3.org/2005/Incubator/ld/XGR-ld-vocabdataset-20111025/>>

□ Collegamento di dati bibliografici con KOSs (*Knowledge Organisation Systems*) e/o altri *dataset* LOD sul Web Semantico (es. contenuti da *DBpedia*, *Open Library*, *Worldcat*) tramite il tool SILK: <<http://wifo5-03.informatik.uni-mannheim.de/bizer/silk/>>

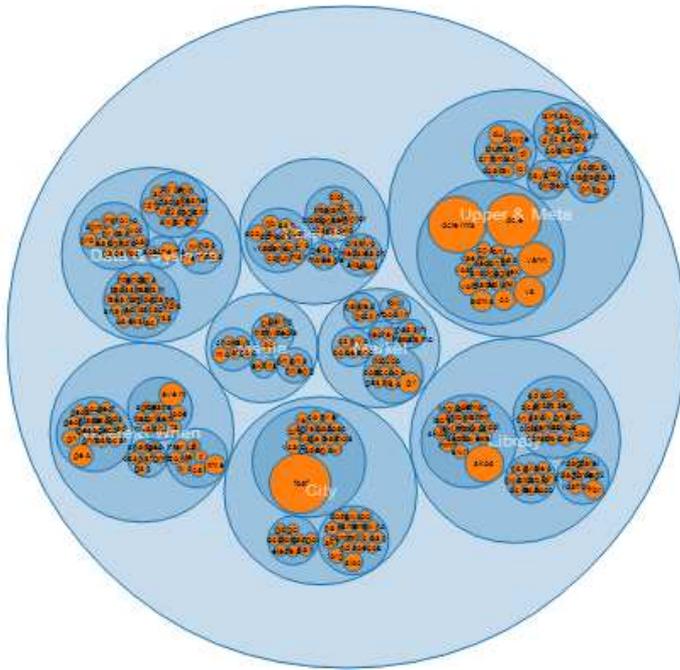
Un caso d'uso: la *Public Library of Veroia* che pubblica i propri contenuti sul Web Semantico/*Web 3.0 (Linked Data Cloud)*:<<http://gr.okfn.org/2012/10/libver/?lang=en>>

***Linked Open Vocabularies (LOV)*:** <<http://lov.okfn.org/dataset/lov/index.html>>

□ È un servizio online (Figura 37) per promuovere e favorire il riutilizzo della straordinaria base di conoscenza fornita dai *Vocabolari Controllati e Semantici* (espressi in RDFS o ontologie OWL) pubblicati come LOD.

□ I 325 vocabolari rappresentati dal servizio vengono:

- ✓ raggruppati nelle seguenti Classi: "*Data & Systems*"; "*Where & When*"; "*City*"; "*Media*"; "*Science*"; "*Market*"; "*Upper&Meta*"; "*Library*";
- ✓ descritti dai metadati;
- ✓ interlineati tra di loro ed espressi nello spazio dei dati interconnessi *Linked Open Data Cloud* - per mezzo il vocabolario VOAF (*Vocabulary of a Friend*, <<http://lov.okfn.org/vocab/voaf/v2.3/>>)



Linked Open Vocabularies (LOV)

Relazioni create tra "Persona", "Organizzazione", "Conferenza".
 Le entità sono definite attraverso *Linked Data*.
Unique Resource Identifier, *URI*, *http* e *Resource Description Framework*, *RDF*.
 I dati pubblicati attraverso *Linked Data* e *Licenze Aperte* diventano *Linked Open Data* ed entrano a far parte dello spazio globale di Web di Dati collegati - *Linked Open Data Cloud*.

Metadata:

Property	Value
Creator	Bernard Vatant, Pierre-Yves Vandenbussche
Contributor	Ghislain Atezeming, Maria Povoda-Villalón
Publisher	Mondeca

```
<?xml version="1.0" encoding="UTF-8"?>
<!-- Generated by RdfSerializer.php from RDF RAP. # http://www.wiwiss.fu-berlin.de/suh/bizer/rdfapi/
<rdf:RDF xmlns:ns3="http://swrc.ontoware.org/ontology#" xmlns:ns2="http://data.semantic
xmlns:ns1="http://xmlns.com/foaf/0.1/" xmlns:vcard="http://www.w3.org/2001/vcard-rdf/
xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:fn="http://www.w3.org/2005/xpath-fu
xmlns:xsd="http://www.w3.org/2001/XMLSchema#" xmlns:owl="http://www.w3.org/2002/
xmlns:sesame="http://www.openrdf.org/schema/sesame#" xmlns:rdfs="http://www.w3.org/
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
- <ns1:Person rdf:about="http://data.semanticweb.org/person/bernard-vatant">
<ns2:holdsRole rdf:resource="http://events.linkedata.org/ldow2010/#pcmemberrole"
<ns1:name xml:lang="en"> Bernard Vatant</ns1:name>
<rdf:type rdf:resource="http://xmlns.com/foaf/0.1/Person"/>
<ns3:affiliation rdf:resource="http://data.semanticweb.org/organization/mondeca"/>
<rdfs:label> Bernard Vatant</rdfs:label>
<ns1:homepage rdf:resource="http://www.mondeca.com"/>
<ns1:made rdf:resource="http://data.semanticweb.org/conference/iswc/2011/paper/
</rdf:Description>
</rdf:RDF>
```

About: [Mondeca](#)
 An Entity of Type : *S.A. (corporation)*, from Named Graph : <http://dbpedia.org>, within Data Space : *dbpedia.org*

DBpedia

Mondeca is a French software company based in Paris. The company name is derived from a place called Mont des Cats, seat of a Cistercian abbey located in famous for its cheese production. The name is a cryptic reference to Marguerite Yourcenar's universe, and to a region with a long tradition of publishing, learning freedom. It has been financed by its founders as well as investment funds Trinova and Banque Populaire. The company currently has 23 employees

Property	Value
dbpedia-owl:abstract	Mondeca is a French software company based in Paris. The company name is derived from a place called Mont

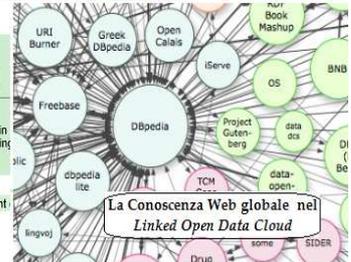


Figura 37. Pagina online del Servizio *Linked Open Vocabularies* (LOV)

□ I vocabolari interconnessi nello spazio del Web di Dati (espressione pratica per eccellenza del Web Semantico) possono stimolare la creatività degli sviluppatori di Archivi digitali nell'originare *nuovi sistemi di navigazione e ricerca* tramite *link a faccette* che esprimono le gerarchie di voci dei vocabolari per migliorare l'espressività dei contenuti depositati.

□ Un esempio di applicazione più interessante, vista la sensibilità verso la tematica e gli importanti costi sostenuti dalle Biblioteche, è quello della *Ricerca Federata* sui Cataloghi delle Biblioteche e sulle Banche Dati (pubblicate in un formato Standard condiviso) esterne fornite dai *Content Provider* internazionali [V. sito: *RDF & Open Linked Data*].

“Your TERMINOLOGY as a part of the SEMANTIC WEB recommendations for design and management” (<<http://www.linkedheritage.org/getFile.php?id=244>>

Traduzione francese: ***Votre terminologie comme élément du web sémantique: Recommandations de conception et de management***”.

Progetto “LINKED HERITAGE”.

I contenuti del documento sono rivolti ad esperti e/o non esperti in *Ingegneria dell'informazione e/o Linguistica*, che hanno un interesse per la Terminologia e il multilinguismo.

Il documento fornisce le raccomandazioni per la gestione delle Terminologie, al fine di renderle compatibili e ottimizzate con i requisiti del multilinguismo e del Web Semantico, per migliorare la ricerca e il recupero delle collezioni digitali *on-line*.

UNBIS THESAURUS: <<http://lib-thesaurus.un.org/LIB/DHLUNBISThesaurus.nsf>>

Il Thesaurus multilingue (Arabo, Cinese, Inglese, Francese, Russo, Spagnolo) multidisciplinare viene impiegato per il controllo semantico d'autorità nel Sistema dell'informazione bibliografica UNBIS delle Nazioni Unite, in particolare per la qualificazione dei valori del SOGGETTO (tema) nei documenti rilevanti per programmi e attività delle Nazioni Unite, essendo stato incorporato nel loro Sistema Documentale Ufficiale.

Consortium of European Research Libraries (CERL) THESAURUS:

<<http://thesaurus.cerl.org/cgi-bin/search.pl?lang=de>>

Il File del Thesaurus CERL è una struttura unica sviluppata per risolvere il problema dell'*impostazione delle varianti* (ortografiche, forme in lingue latine e altri, nomi fittizi) per i *nomi* (personali, geografici, della misurazione ufficiale) usati nell'Europa agli inizi del periodo di stampa (tra 1450 - e 1830 d.C.). Come tale, è uno strumento di ricerca indispensabile per studiosi e ricercatori di questo periodo.

European Language Social Science THESAURUS (ELSST):

<<http://elsst.esds.ac.uk/Home.aspx>>

È un Thesaurus multilingue (Inglese, Danese, Finlandese, Francese, Tedesco, Greco, Norvegese, Spagnolo, Svedese e Tedesco) per l'uso in un contesto internazionale e multidisciplinare per le SCIENZE SOCIALI (la politica, la sociologia, l'economia, l'istruzione, la legge, la criminalità, la demografia, la salute, l'occupazione, la tecnologia).

International Labour Organization (ILO) THESAURUS:

<<http://www.ilo.org/public/english/support/lib/tools/aboutthes.htm>>

È un Thesaurus multilingue (Inglese, Francese, Spagnolo), contiene termini su una vasta gamma di argomenti legati allo SVILUPPO ECONOMICO e sociale. Le aree di maggior interesse sono trattate in dettaglio. Esse includono: la politica del lavoro e dell'occupazione, le norme sul lavoro, amministrazione del lavoro, rapporti di lavoro, la formazione professionale, la protezione sociale, condizioni di lavoro, i salari, la sicurezza e salute sul lavoro, piccola impresa, l'economia informale e dei diritti umani sul posto di lavoro.

National AGRICULTURAL Library Thesaurus (NAL) THESAURUS:

<<http://agclass.nal.usda.gov/download.shtml>>

Il Thesaurus è organizzato in 17 categorie di soggetti per facilitare la navigazione tramite le gerarchie molto granulari.

UNESCO THESAURUS: <<http://databases.unesco.org/thesaurus/>>;

<<http://skos.um.es/unescothes>>

È un Thesaurus multilingue (Inglese, Russo, Francese, Spagnolo) disponibile online come *Linked Open Data* anche con i formalismi del Web Semantico SKOS (*Simple Knowledge Organization System*).

Rappresenta una lista gerarchica di termini controllati utilizzati per l'analisi del SOGGETTO dei documenti e per la ricerca del materiale in materia di *istruzione, cultura,*

scienze naturali, scienze umane e sociali, diritto, comunicazione e informazione.

È strutturato in 7 aree (domini) disciplinari principali suddivisi in micro-Thesauri che permettono di ottenere una rapida panoramica delle unità concettuali.

EUROVOC THESAURUS: <<http://eurovoc.europa.eu/drupal/>>

È un Thesaurus multilingue (24 lingue) che copre la descrizione del SOGGETTO nelle attività dell'Unione Europea e del Parlamento europeo.

AGROVOC THESAURUS:

<<http://aims.fao.org/standards/agrovoc/linked-open-data>>;

<<http://aims.fao.org/standards/agrovoc/about>>

Contiene 32.188 concetti (in 22 lingue) del campo di agricoltura.

Disponibile in modalità *Linked Open Data* (LOD) e tramite i formalismi SKOS (*Simple Knowledge Organization System*).

THESAURI e vocabolari del patrimonio culturale disponibili come *Linked Open Data*:

<<http://www.heritagedata.org/blog/about-heritage-data/seneschal/>>.

SENESCHAL project: <<http://www.heritagedata.org>>

Thesauri e vocabolari controllati che includono concetti relativi ai *tipi di monumenti, reperti archeologici, eventi, periodi di tempo* e le *imbarcazioni marittime* sono disponibili anche per il *download* in formato SKOS.

ISO 25964-2 (2013): *Information and documentation -- THESAURI and interoperability with other vocabularies*: <<http://nkos.slis.kent.edu/PressReleaseISO25964-2.pdf>>

È uno Standard che tratta la questione dell'interoperabilità tra *Thesauri* e altri vocabolari (schemi di Classificazione, Tassonomie, Schemi di intestazione di soggetto, Liste di autorità dei nomi, Ontologie, Terminologie).

L'interoperabilità è intesa come la capacità di due o più sistemi o componenti di scambiare informazioni, e utilizzare le informazioni scambiate.

Standard tratta i seguenti argomenti: modelli strutturali e Linee Guida per la mappatura (*crosswalking*) tra *Thesauri* e vocabolari, la gestione *pre-coordinata dei concetti* che si verifica soprattutto in schemi di Classificazione, Tassonomie e Schemi di intestazione di soggetto.

GeoNames ONTOLOGY:

<<http://www.geonames.org/ontology/documentation.html>>; <<http://www.geonames.org/>>

L'Archivio ovvero l'Ontologia geografica più grande al mondo disponibile in modalità LOD che contiene i dati principali di tutte le città e nazioni, tutti i laghi, le montagne più alte e altri milioni di *dati* (descrizioni RDF più di 8.000.000 nomi) *geografici a livello globale*.

È un progetto dove chiunque può contribuire inserendo o correggendo toponimi che possono essere liberamente utilizzati attraverso servizi Web nella la ricerca (tramite flussi RSS) nella descrizione nei dei nomi delle città con la successiva associazione ad ogni elemento delle coordinate geografiche del luogo citato.

Bibliographic ONTOLOGY (BIBO): <<http://bibliontology.com/>>

È un'Ontologia per descrivere le *entità bibliografiche* sul Web Semantico attraverso i linguaggi ontologici web OWL/RDFS.

Bibliographic Ontology Specification: <<http://bibliontology.com/specification>>

Le specifiche BIBO forniscono concetti e proprietà fondamentali per *descrivere citazioni e riferimenti bibliografici*, ad esempio, di libri, articoli nel Web Semantico.

FRBR –aligned Bibliographic Ontology (FaBiO):

<<http://www.essepuntato.it/lode/http://purl.org/spar/fabio>>

È un'Ontologia per la registrazione e la pubblicazione sul Web Semantico di *record*

bibliografici prodotti nel campo di *ricerca scientifica* (libri, riviste, e oggetti del loro contenuto, come articoli, parti del libro, post di conferenza ecc.)

□ Le classi di FaBiO sono strutturati secondo lo schema FRBR:

“Opere”, “Espressioni”, “Manifestazioni”, “Entita”.

□ Le proprietà aggiuntive servono per estendere il modello di dati FRBR mediante *linking* di

✓ “Opere” e “Manifestazioni” (es. “*fabio:hasManifestation*”

e “*fabio:isManifestationOf*”);

✓ “Opere” e “Entita” (es. “*fabio:hasPortrayal*” e “*fabio:isPortrayedBy*”);

✓ “Espressioni” e “Entita” (es. “*fabio:hasRepresentation*” e “*fabio:isRepresentedBy*”).

Friend of a Friend (FOAF) ONTOLOGY: <<http://www.foaf-project.org/>>

□ È un' Ontologia Web che, basandosi sui linguaggi Web ontologici RDF e OWL, descrive gruppi di persone e le loro relazioni

(es. “*Friend Of*”, “*Son Of*”, “*Lives In*”, “*Works In*”, “*Knows Someone*”, “*Foaf:Page*”, “*Foaf:Focus*”, “*Foaf:isPrimaryTopicOf*”).

□ Il Catalogo digitale “*Open Library*”: <<http://openlibrary.org/>> (Figura 38) è un esempio d'eccellenza di implementazione dell'Ontologia FOAF.



Figura 38. Pagina online del Catalogo digitale “*OPEN LIBRARY*”

Nello *screenshot* rappresentato nella Figura - che ha catturato i risultati tramite la voce di ricerca “*Ferdinand de Saussure*” - vengono elencati i libri di *FERDINAND DE SAUSSURE* e le categorie “SUBJECTS”, “PLACES”, “PEOPLE”, “TIMES”, “PROLIFIC AUTHORS”, “PROMINENT PUBLISHERS” collegati all’oggetto di ricerca tramite l’Ontologia FOAF.

La piattaforma “*OPEN LIBRARY*” persegue l’obiettivo di mettere a disposizione *una pagina web* per *ogni libro* pubblicato nel mondo, conservato presso una Biblioteca o venduto su Internet.

“*OPEN LIBRARY*” presenta 23 milioni di opere (di cui 1 milione liberamente scaricabili), incluso circa 900.000 di testi digitalizzati provenienti da *Google Books Search*).

ONTOLOGY plug-in per l’integrazione di vocabolari controllati nei sistemi gestionali dell’informazione (information management systems):

<<http://aims.fao.org/ontology-plugin-webinars-at-aims>>

“*Ontology plug-in*”⁸ per la ricerca sul web e l’integrazione dei vocabolari controllati nei sistemi gestionali dell’informazione: <<http://www.youtube.com/watch?v=VHAMfk-q0Mo>>; <<http://www.slideshare.net/faoaims/the-new-ontology-plugin>>

Ontology plug-in cerca contemporaneamente multipli Thesauri e Ontologie utilizzando un servizio web broker come AGROVOC, ASFA, Ontologia Plant, NERC-C19 ontology, e OceanExpert, fornendo un elenco di concetti selezionati identificati con URIs e le etichette preferite.

Pur essendo sviluppato nel contesto dell’iniziativa AgriOcean DSpace (<<http://aims.fao.org/tools/agrioccean-dspace>>), *Ontology plug-in* può essere utilizzabile in qualsiasi altra applicazione del sistema informatico (es. *Repository*, Biblioteca digitale).

CONCEPT MAPPER.

Strumento di analisi dei testi e di generazione automatica di metadati semantici.

CINECA for the Digital University, 2013:

<http://www.cineca.it/sites/default/files/Cineca_DigitalUniversity.pdf>

E’ un *use case* dello strumento di analisi dei testi integrato tramite API nei seguenti prodotti del CINECA:

- ✓ *Sapienza Digital Library Management System*;
- ✓ *L2L (Live to Learning)*;
- ✓ *Navigazione Concettuale dell’offerta didattica (EduMap)*.

Sfrutta la conoscenza disponibile in diverse risorse linguistiche (Wikipedia, EuroVoc, altre ontologie, dizionari e Thesaurus specialistici) per:

Identifica (anche tramite le trascrizioni del parlato) ed estrapola i concetti più rilevanti nel contesto del documento (quelli che maggiormente riassumono il suo significato) e/o in relazione al dominio di riferimento, tramite il calcolo delle misure di vicinanza semantica, associando le parti del testo a tutti gli altri concetti individuati nel contesto del documento (valori di *internal relatedness*) e in altre fonti (valori di *external/domain relatedness*).

Alti valori di *internal* e *external relatedness* (la loro somma deve essere maggiore a 0,5) è un criterio sufficiente per operare una prima selezione dei concetti. I concetti così individuati vanno ad arricchire i metadati descrittivi già associate alle entità digitali, permettendo di offrire all’utente una nuova “dimensione” (*facet*) durante la fase di *browsing*.

Esegue la disambiguazione (identificazione del concetto in maniera univoca), validazione, filtro dei concetti;

Esegue l’annotazione automatica di parti del testo con i concetti corrispondenti ed eventualmente con il link alla pagina della fonte che ne fornisce la definizione e la descrizione;

Traduce i concetti in altre lingue, o altre informazioni disponibili nelle fonti utilizzate;

Esegue la mappatura del contenuto del documento su un’Ontologia specifica. Associa metadati semantici quali le categorie di appartenenza dei concetti e le loro relazioni, per creare il contesto di riferimento di ciascun concetto dell’ontologia (gruppo di concetti correlati, “*group relatedness*” calcolata sulla base dei link condivisi dalle rispettive pagine che descrivono il concetto).

Al termine dell’elaborazione della mappatura viene generato un file RDF SKOS (all’interno con i concetti formati da metadati e concetti associati) che può essere inserito in un *datastream* del *Software Fedora* e indicizzato per la ricerca dei contenuti.

⁸ *PLUG-IN* e’ un programma non autonomo che interagisce con un altro programma per ampliarne le funzioni. La capacità di un Software di supportare i plugin e’ generalmente un’ottima caratteristica, perché rende possibile l’ampliamento e la personalizzazione delle sue funzioni in maniera semplice e veloce.

Il formato SKOS utilizzerà come identificativo l'identificatore web URI del dataset DBPedia per descrivere il concetto, allo scopo sia di indicizzazione che di collegamento con i *Linked Data*.

□ Il mapping all'Ontologia viene sfruttato all'interno della "*Cross Discipline Search*".

□ I metadati vengono utilizzati dal prototipo all'interno della "*SEMANTIC SEARCH*", che, a differenza della "*Keyword Search*", aggiunge alla ricerca full-text la ricerca tra i sinonimi e le categorie associate al testo (metadati).

□ *CONCEPT MAPPER* è di facile integrazione e può essere adoperato da altre applicazioni per:

- ✓ raggruppare i documenti in categorie significative;
- ✓ classificare i documenti in categorie predefinite;
- ✓ realizzare mappe di navigazione concettuale;
- ✓ espandere i metadata attraverso i *Linked Data*;
- ✓ indicizzare semanticamente contenuti da parte dei motori di ricerca [Paci et al., 2010].

Le funzionalità di *Concept Mapper* si basano sui processi di *Knowledge Discovery e Management* rappresentati nella Figura 35.

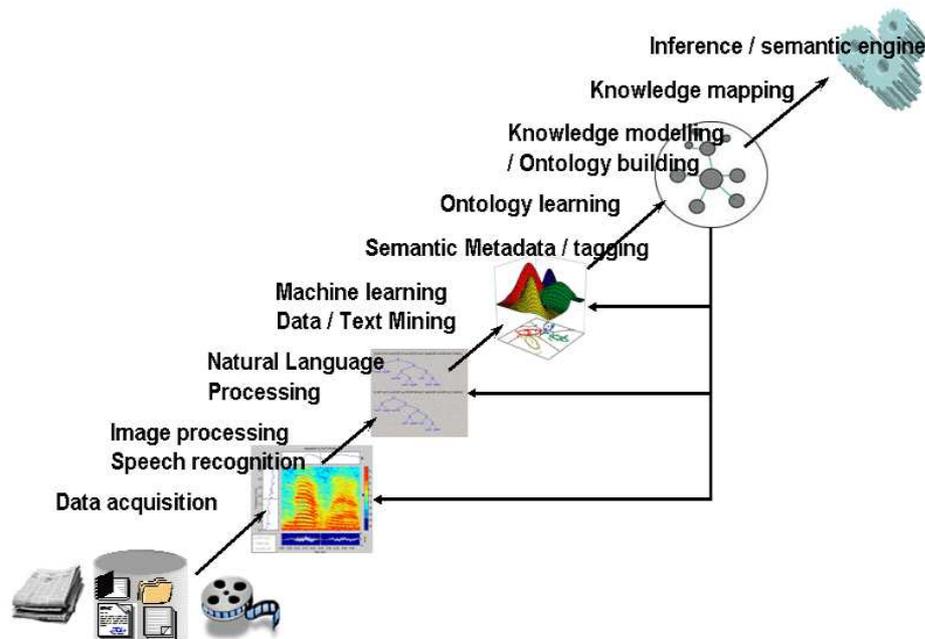


Figura 39. Processi di *Knowledge Discovery e Management* attivati dal tool "*Concept Mapper*"

THESAURUS-augmented Search with Jena Text

<http://zbw.eu/labs/en/blog/thesaurus-augmented-search-with-jena-text>

La ricerca tramite l'implementazione del "*STW Thesaurus for Economics*" i cui concetti sono mappati con "*Thesaurus for the Social Sciences*" (Figura 40).

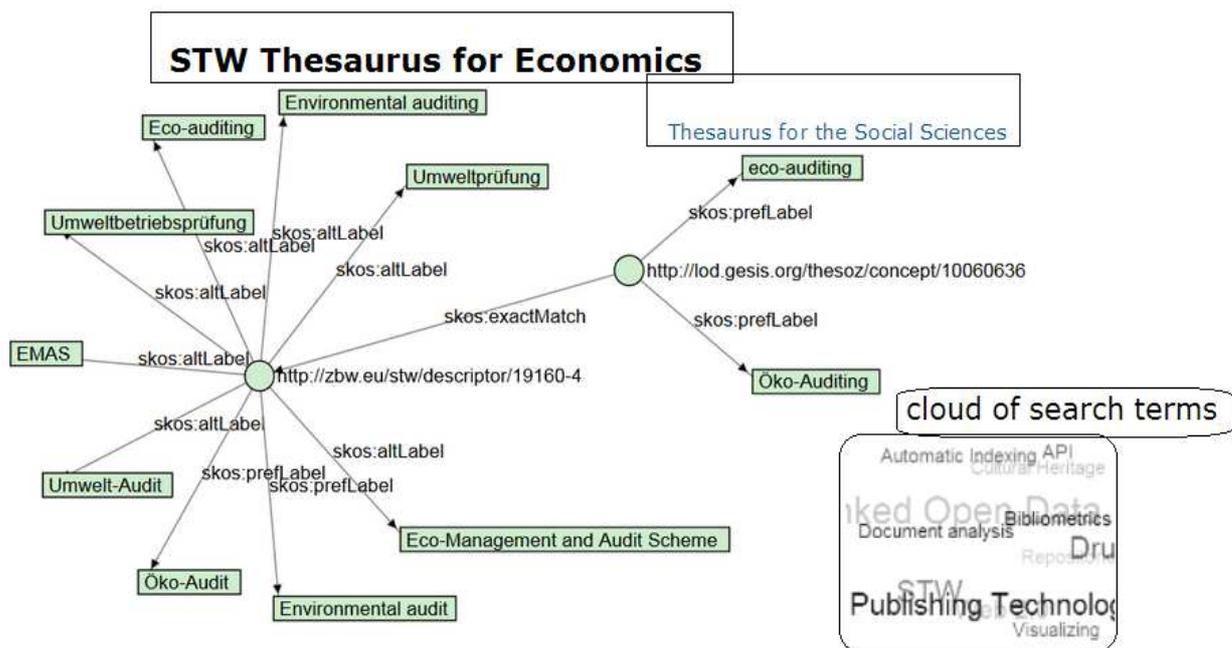


Figura 40. La mappatura del “STW Thesaurus for Economics” con il “Thesaurus for the Social Sciences” per definire i termini per la ricerca nella *Cloud of search terms*

L'integrazione del Thesaurus STW può essere fatta sul “lato server” (*server-side*), nelle applicazioni *backend*. Oppure può essere realizzata sul “lato client” (*client-side*) trascinando insieme i sinonimi in una stringa di ricerca (per Google, questo potrebbe essere rappresentato da "un termine A" OR "termine B" OR ...), mettendola nella casella (box) di ricerca, e avviando la procedura di ricerca attraverso Javascript.

***The International Society for Knowledge Organisation (ISKO):* <<http://www.iskouk.org/>>**

Dei KOS e degli approcci semantici all'informazione si occupa un'associazione internazionale *International Society for Knowledge Organization (ISKO)*.

ISKO Italia è un capitolo nazionale dell'associazione: <<http://www.iskoi.org/>>

Nonostante gli indiscutibili vantaggi della qualificazione dei valori metadati tramite le entità semantiche offerte da diversi KOS, questo processo potrebbe rappresentare anche uno degli aspetti più complessi dell'organizzazione dei contenuti in un Archivio Istituzionale, in quanto le varie Classificazioni, Thesauri o Terminologie strutturate, pur interessando la *stessa area*, possono mantenere

forti disaccordi sul piano linguistico, strutturale e semantico, con divergenze particolarmente gravi tra le Classificazioni disciplinari e le Classificazioni generali, quali la Classificazione Decimale Dewey, la Classificazione Decimale Universale, o la Classificazione della Library of Congress [Guerrini, 2010a, p.49].

Per contrastare tali problematiche, occorre contribuire allo sviluppo e all'implementazione:

- ✓ dei *Thesauri specializzati* per gli Archivi Istituzionali. Questo è possibile con l'impiego delle *Liste di Autorità* con cui *organizzare la conoscenza* di un *numero limitato di termini/collezioni lessicali* (sia di significato generale che specifico) individuati e ricavati dal *corpus di metadati o documenti a testo completo* contenuti nell'Archivio.

Affinché i Thesauri sviluppati siano facilmente integrabili nel Web semantico, condivisi e utilizzabili dai diversi agenti della rete, è necessario che si adottino i formalismi del Web Semantico *Simple Knowledge Organization System* (SKOS) e che si pubblichino come Linked Open Data.

- ✓ delle *equivalenze linguistiche* [V. sito: “L’equivalenza nella terminologia multilingue“,] in diverse lingue per la terminologia contenuta nei medesimi *Thesauri*.

4.5. Conservazione a lungo termine

La *qualità* di un IR non può prescindere da una *corretta* impostazione delle *Politiche di conservazione* (*Digital Preservation*), cioè dall’*insieme* delle *azioni* e degli *interventi* necessari a garantire un *costante* e *affidabile accesso* agli oggetti digitali per tutto l’arco di tempo in cui vengono considerati utili. La Conservazione delle risorse digitali è un problema da affrontare con attenzione e urgenza, pena una perdita globale della memoria di dimensioni finora sconosciute [...]

Troppe istituzioni, nella fase di programmazione di un IR, tengono poco conto delle problematiche legate alla Conservazione. La programmazione di periodici *backup* dei dati *non* può essere ritenuta una *misura sufficiente* a garantirne la fruibilità nel lungo periodo [...] Vanno favoriti, infine, *l’aggiornamento continuo del personale*, la *collaborazione tra professionisti ed esperti* di vari ambiti disciplinari e il coinvolgimento degli autori nelle *Politiche di conservazione dei contenuti digitali* [Guerrini, 2010a, pp. 57-58, 60].

In futuro, una parte consistente – probabilmente la più rilevante – della *memoria delle Istituzioni scientifiche, accademiche e culturali* sarà formata da oggetti “*nati*” *digitali*, conservati in Archivi ad Accesso Aperto. Per garantire la *sopravvivenza* di queste *risorse* e la loro *citazione*:

- ✓ occorre che le comunità accademiche e scientifiche si impegnino fortemente nella *conservazione a lungo termine* dei propri Archivi digitali [...]
- ✓ L’ACCESSO deve essere GARANTITO, fintanto che è ritenuto *necessario*, *prevenendo* e *controllando* le *alterazioni* ed i guasti dei supporti e dei sistemi, le anomalie di funzionamento dei *Repository* e gli effetti dell’*evoluzione tecnologica*. A tal fine
- ✓ i processi per la CONSERVAZIONE A LUNGO TERMINE devono essere basati su Modelli organizzativi, Politiche e Strategie ben definite [...]

L’AFFIDABILITÀ si riferisce alla capacità delle componenti *Hardware* e *Software* di operare e comportarsi in base alle specifiche, senza anomalie o errori (in teoria completamente, in pratica garantendo una percentuale molto elevata). La DISPONIBILITÀ è la percentuale di tempo nella quale un sistema, un’applicazione o un componente funzionano regolarmente rispetto al tempo totale di funzionamento richiesto. I *salvataggi* di dati e delle applicazioni, gli *antivirus*, i *firewall*, gli *aggiornamenti* dei sistemi operativi e dei Software applicativi, le *configurazioni Hardware* basate su componenti *fault tolerant* e ridondanti sono alcune delle tecniche comunemente adottate per garantire elevate percentuali di affidabilità e disponibilità.

[Pasqui, 2012, pp.1-2]

L’Archivio digitale di qualità deve garantire l’accesso e la conservazione a lungo termine dei contenuti che ospita.

l’espressione italiana “*Conservazione digitale*” indica tendenzialmente il processo di digitalizzazione di documenti analogici, mentre “*Conservazione del documento/memoria/informazione/risorsa digitale*” è la più comune traduzione dell’espressione *DIGITAL PRESERVATION* [...] si tenda a sottovalutare la valenza della coppia terminologica *Conservation/Preservation*.

Mantenere anche in italiano la distinzione tra i due termini potrebbe essere funzionale per segnalare un *binomio semantico* che ha i suoi corrispettivi sia nell’ambito delle modalità di *trasmissione dell’informazione(analogico/digitale)* che di tipologia dell’intervento (*curare/prevenire*); quindi “PRESERVAZIONE” indicherebbe meglio un’attenzione funzionale non solo a *contenere* ma a *prevenire gli effetti dell’obsolescenza* e della fragilità in contesti digitali

[Faiella, 2000, p.64]

Enormi quantità di informazioni tra loro molto differenti sono oggetto di *codifica digitale*; alcune tipologie di dati sono assimilabili ai documenti cartacei – ad esempio le biblioteche conservano documenti stampati sugli scaffali dai quali in futuro potranno essere prelevati e letti. Una situazione analoga a *livello digitale* significa che, *in futuro*, dovrà essere possibile *prendere un file COMPUTERIZZATO e STAMPARLO* [V. sito: Ricerca nel campo delle TIC].

Il tema della conservazione a lungo termine di oggetti digitali dovrebbe essere messo al centro già nella fase di progettazione del sistema Archivio, considerando che senza una preventiva programmazione degli aspetti della preservazione (“*Preservation Planning*”) dei dati elettronici racchiusi in diverse tipologie di documenti (compresi quelli dei risultati scientifici⁹), una grande massa di tali dati col tempo potrà rivelarsi illeggibile o a rischio, perché le nuove tecnologie non ne consentiranno la lettura o perché risulteranno incomprensibili agli utenti.

Secondo lo Standard per l’archiviazione e la conservazione delle risorse digitali a lungo termine - ISO 14721:2012 OAIS (*Open Archival Information System*), la RISORSA DIGITALE è un “*INFORMATION PACKAGE*” (OAIS-IP) composto da:

1. METADATI che descrivono una RISORSA INFORMATIVA dal punto di vista del contenuto, amministrativo, strutturale, di provenienza;
2. OGGETTI DIGITALI ovvero i *file*, compreso i *file* multimediali che compongono la risorsa.

L’OAIS-IP può riferirsi ad altri due tipi di entità:

1) la COLLEZIONE ORGANIZZATIVA (un insieme di risorse digitali gestite e/o prodotte nell’ambito di competenza di una qualsiasi struttura organizzativa dell’Ateneo: Biblioteca, Dipartimento, Centro, Istituto ecc.).

Le *Collezioni Organizzative non* possono essere *strutturate gerarchicamente* ma si presentano in *forma atomistica* nei confronti dell’Ente che le ospita.

Il materiale di una Collezione Organizzativa può:

- ✓ contenere al suo interno sia le Risorse digitali singole che
 - ✓ organizzato e raccolto in *Collezioni di Risorse*, sulla base di un legame logico, tematico, gestionale o tramite una relazione di altra natura;
- 2) la COLLEZIONE di RISORSE:
- ✓ è associata necessariamente ad una *Collezione Organizzativa*;
 - ✓ può contenere zero o più sotto-collezioni di risorse.

Il PACCHETTO INFORMATIVO “OAIS-IP” viene rappresentato da un *file* basato su uno schema di metadati METS (*Metadata Encoding and Transmission Standard*) in cui vengono descritti tutti i componenti costitutivi di una risorsa incluso i metadati. Le informazioni che secondo la terminologia OAIS possono essere considerate il PDI (*Preservation Description Information*) sono anch’esse contenute nel *file* METS.

Prima di implementare un’infrastruttura che possa ospitare le risorse digitali, occorre:

- considerare le prestazioni del *Software* scelto in relazione a:
 - ✓ capacità di immagazzinamento dei contenuti digitali;
 - ✓ capacità di adeguarsi alle misure per garantire l’accesso agli oggetti a lungo termine;
 - ✓ supporto dei formati di oggetti digitali per la conservazione a lungo termine;
 - ✓ capacità per l’assegnazione e la gestione di identificatori persistenti. Bisogna prevenire il fenomeno dei link interrotti (il ben noto errore “*404 Not Found*”) adottando soluzioni tecniche ed organizzative che superino i limiti delle citazioni basate sugli URL non persistenti;
 - ✓ possibilità di migrazione dei dati da una piattaforma ad un’altra;
 - ✓ meccanismi di manutenzione, *backup*, *disaster recovery* e sicurezza;
- valutare quali Standard possano essere applicati per mantenere inalterati (attendibili) file nel tempo;
- abbozzare le Politiche organizzative e Strategie per valutare la fattibilità del sistema in vista della persistenza e della conservazione degli oggetti digitali, e per definire gli ASPETTI

⁹ Il sito della Community CASPAR (<<http://www.casparpreserves.eu/>>) offre una serie di video istruttivi a supporto delle tematiche di preservazione digitale. Il video “*ESA-ESRIN: Preserving Digital Data within the Scientific Domain*” parla dell’importanza di preservazione di dati scientifici dell’*European Space Agency* e dell’*Advanced Computer Systems*.

ORGANIZZATIVI compreso i RUOLI che possono essere richiesti a supporto dei processi di conservazione.

In un Archivio implementato è necessaria una costante revisione e valutazione (*auditing and assessment*) delle Politiche e Strategie definite a supporto dell'intero ciclo di vita di oggetti digitali, per poter tempestivamente far fronte ai cambiamenti di tipo organizzativo, delle Politiche legali e tecnici.

In seguito saranno esposte informazioni, elencate fonti, e forniti rimandi di approfondimento relativamente al tema del presente paragrafo.

Open Archival Information Systems (OAIS). Standard ISO 14721:2012
Reference Model for OAIS: Recommendation for Space Data System Practices.
Versione integrale: <<http://public.ccsds.org/publications/archive/650x0m2.pdf>>

□ OAIS è “an ARCHIVE, consisting of an organization of People and Systems, that has accepted the responsibility to Preserve Information and make it available for a Designated Community [...]

□ The term "OPEN" in OAIS is used to imply that ISO 14721:2012, as well as future related International Standards, are developed in OPEN FORUMS, and it does not imply that access to the archive is unrestricted.

I principi definiti dall'OAIS Reference Model (Figura 41) è un prerequisito chiave per la costruzione di Repository affidabili (Trusted) e per assicurare la conservazione a lungo termine di risorse informative che essi ospitano.

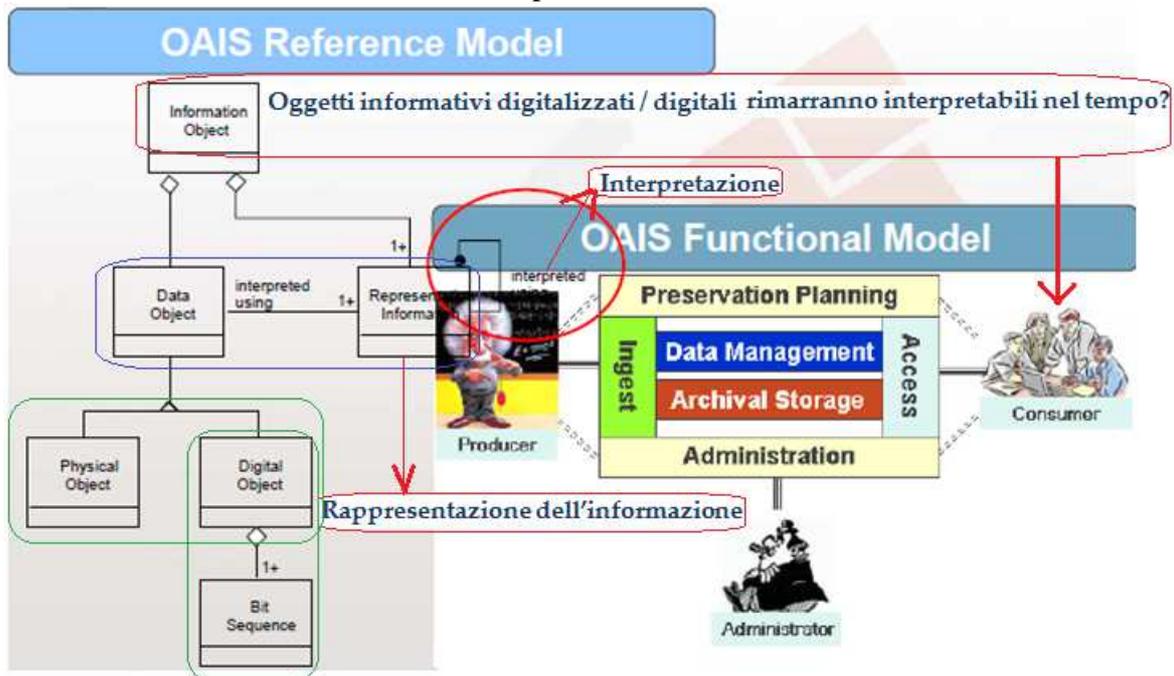


Figura 41. Open Archival Information System (OAIS) - Reference model

OAIS definisce un pacchetto di informazioni d'ARCHIVIO (*Archival Information Package*) che contiene tutte le informazioni necessarie per la Conservazione a lungo termine di un oggetto informativo digitale.

Confronto fra Software per la gestione di Digital Repository (DSpace, Fedora, Digitool, aDORe, Greenstore, MyCoRe). Progetto Magazzini Digitali:

<<http://www.rinascimento-digitale.it/documenti/confrontorepository/ConfrontoRepository2006.pdf>>

Tra i criteri di inclusione nell'analisi dei *Software* ci sono:

- il supporto per lo Standard ISO 14721:2012 OAIS (*Open Archival Information System*) per la conservazione delle risorse digitali a lungo termine;
- il supporto per gli identificatori persistenti *URL*, importanti nell'ottica della garanzia dell'accesso ai dati digitali a lungo termine.

“Digital preservation & records management Programme”.

JISC (Joint Information Systems Committee):

<http://www.jisc.ac.uk/whatwedo/programmes/preservation.aspx>

- La conservazione dei contenuti digitali è una serie di attività gestionali necessarie a garantire un accesso significativo a lungo termine agli oggetti digitali.
- La necessità di sviluppo di un *Information Environment* volto a supportare gli Archivi digitali, la Conservazione del digitale, *cross-searching* tra gli Archivi.

JISC “Digital Preservation briefing paper”

http://www.jisc.ac.uk/publications/briefingpapers/2006/pub_digipreservationbp.aspx

- Continued access to authentic digital assets.** Gli assetti digitali (es. documento, collezione, *ebook*, aggregazione, immagine, mappa, video) di un Archivio devono essere preservati per garantire la loro autenticità per il futuro utilizzo.
- What is Digital Preservation?** La preservazione digitale è una serie di azioni e di interventi necessari per assicurare l'accesso continuo e affidabile a oggetti digitali autentici;
- Digital Curation.** La gestione attiva e la valutazione dei contenuti digitali attraverso il monitoraggio del loro ciclo di vita. *Digital Curation* si basa sui concetti di base della conservazione digitale, sottolineando le opportunità di valore aggiunto e della conoscenza attraverso l'annotazione e la continuità nella gestione delle risorse;
- Digital Curation Centre (DCC).** Fornisce un piano per la ricerca e lo sviluppo delle questioni sulla *Digital Curation* e promuove esperienze e buone pratiche per la gestione dei risultati di ricerca in formato digitale.

DCC Curation Lifecycle Model: <http://ijdc.net/index.php/ijdc/article/view/69>

Tale Modello può essere utilizzato in combinazione con gli Standard rilevanti per pianificare le attività di Conservazione a lungo termine a diversi livelli. Le strategie di Conservazione devono essere messe in atto sia a livello degli Archivi (*Repository*) che degli oggetti digitali.

PREMIS (PREservation Metadata: Implementation Strategies).

PREMIS Data Dictionary for Preservation Metadata (v. 2)

<http://www.loc.gov/standards/premis/>

PREMIS è uno Standard internazionale sviluppato sul Modello OAIS per:

- identificare e valutare STRATEGIE per la *codifica*, la *memorizzazione* e la gestione dei *metadati* nei sistemi di *Digital Preservation*;
- fornire i metadati per sostenere la Conservazione di oggetti digitali e per garantirli l'usabilità a lungo termine.

Preserving Access to Digital Information (PADI):

<http://pandora.nla.gov.au/pan/10691/20110824-1153/www.nla.gov.au/padi/topics/18.html>

È un sito pubblicato dalla Biblioteca Nazionale di Australia. Offre una guida sintetica alle principali strategie per la Conservazione digitale:

- la migrazione (*migration*): il trasferimento dei dati su un nuovo sistema anche su supporti digitali diversi. Questa strategia può comportare la conversione delle risorse da un formato a un altro;

il *refreshing*: il trasferimento di dati tra due supporti digitali dello stesso tipo; in questo modo non ci sono cambi o alterazioni dei dati;

la duplicazione (*replication*): la creazione di copie sui dati in più sistemi;

l'emulazione (*emulation*): la duplicazione delle funzionalità di un sistema obsoleto. L'emulazione richiede la creazione di emulatori, programmi che traducono il codice e le istruzioni di un ambiente-macchina in modo da farli girare correttamente in un altro;

l'incapsulamento (*encapsulation*) dei dati al fine di garantire la protezione degli stessi; Così come le risorse informative digitali, anche i metadati (informazioni sulla *creazione*, *diritti di accesso*, *restrizioni di accesso*, *storia di preservazione* e *diritti per l'uso*) possono subire l'obsolescenza. Il formato ASCII (*American Standard Code for Information Interchange*), che utilizza caratteri leggibili e non codici numerici e che è largamente diffuso, è considerato il formato più durevole per i metadati.

“Politiche di gestione della Conservazione delle collezioni digitali:

Strategie a breve termine per contrastare problemi a lungo termine”:

<<http://www.dpworkshop.org/dpm-ita/foundation/oais/index.html>>

Questo documento contiene 6 Capitoli e 6 Appendici che forniscono le informazioni fondamentali per la realizzazione di un'infrastruttura digitale conforme allo standard OAIS. Presenta, inoltre, i settori che sono oggetto di studi ulteriori da parte dei gruppi di lavoro OAIS.

Digital Preservation Europe (DPE):

<<http://www.digitalpreservationeurope.eu/registries/resources/>>

È una collezione di risorse *on-line* che riporta le informazioni utili sugli studi e sui progetti di Archivi digitali.

Digital Preservation Coalition (DPC): <<http://www.dpconline.org/>>

È un'organizzazione *non-profit* costituita in Inghilterra per coordinare una serie di azioni per affrontare le sfide derivanti dalla tematica "Conservazione del digitale" e per condividere la medesima conoscenza a livello internazionale.

Open Archives Initiative Object Reuse and Exchange (OAI-ORE)

<<http://www.openarchives.org/ore/>>

OAI-ORE definisce Standard per la descrizione e lo scambio di aggregazioni delle risorse Web. Lo scopo di questi Standard è di esporre i contenuti delle medesime aggregazioni *alle* applicazioni che supportano controllo di autorità, deposito, scambio, visualizzazione, riuso, e preservazione.

“*What Does OAI-ORE mean for Digital Preservation?*”:

<http://www.ariadne.ac.uk/issue57/rumsey-osteen#What_Does_OAI-ORE_Mean_for_Digital_Preservation>

“Linee Guida DRIVER”:

< http://www.driver-support.eu/documents/DRIVER_Guidelines_v2_Final_2008-11-13.pdf> **Chapter 7: Annex: Use of PERSISTENT IDENTIFIERS:**

“*PERSISTENT IDENTIFIERS for web resources are needed to create a stable and reliable infrastructure. This does not concern technicalities, but mainly agreements on an organizational level*”.

“Linee Guida per gli Archivi Istituzionali” (CRUI, Italia):

<<http://www.cruil.it/HomePage.aspx?ref=1781>>.

Sono create in conformità alle “Linee Guida OpenAIRE”

(<<http://www.openaire.eu/it/component/attachments/download/79%E2%8C%A9=en>>) alla loro volta basate sulle “Linee Guida DRIVER”.

Nel paragrafo “PRESERVAZIONE” vengono indicati, in modo sintetico, gli aspetti principali di una POLITICA di CONSERVAZIONE delle risorse digitali, in cui viene

accentuata l'importanza di:

- un'accurata selezione del materiale a cui si deve garantire la persistenza dell'accesso a lungo termine;
- lo sviluppo delle POLITICHE di CONSERVAZIONE in conformità allo Standard per la Conservazione del digitale OAIS (strumenti concettuali e indicazioni organizzative per poter predisporre un'adeguata progettazione degli *interventi conservativi* fin dal momento della creazione dell'Archivio);
- la gestione di IDENTIFICATORI PERSISTENTI (*PERSISTENT IDENTIFIERS*) basati su URIs (*UNIFORM RESOURCE IDENTIFIERS*) forniti di un meccanismo di risoluzione agli identificatori permanenti (Handle, DOI, URN);
- lo sviluppo di sistemi di *Storage - Backup - Business Continuity - Disaster Recovery - reportistica (log)*;
- la gestione dei formati (conversione, migrazione, emulazione);
- la gestione di metadati per la conservazione;
- la garanzia dell'autenticità (integrità, firma digitale) dei dati .

IDENTIFICATORI PERSISTENTI (PID)

“Per risolvere il problema delle versioni e delle copie autentiche si devono individuare *sistemi affidabili di citazione* delle risorse, basati su *Identificatori Unici e Persistenti*”
[Guerrini, 2010a, p.59]

L'associazione di un Identificatore Persistente a una risorsa digitale è considerata oggi una *buona pratica* che contribuisce alla risoluzione del problema del *reperimento affidabile* delle risorse in rete e, allo stesso tempo, può *supportare la certificazione dell'autenticità* della risorsa, della sua *provenienza*, dei suoi diritti di *sfruttamento*
[V. sito: “Deposito Legale“]

URN (*Uniform Resource Name*)

<<http://www.iana.org/assignments/urn-namespaces/urn-namespaces.xml>>

È un URI (*Uniform Resource Identifier*) che identifica una risorsa mediante un "nome" in un particolare dominio di nomi ("*namespace*").

A differenza del URL (*Uniform Resource Locator*), URN non permette l'identificazione della locazione (l'indirizzo) della risorsa in rete.

Un esempio di URN è il codice ISBN: questi identifica univocamente un libro, ma non ci dà alcuna informazione sulla locazione dello stesso.

Un URN deve rimandare almeno a un URL attraverso cui accedere all'oggetto.

L'URN collega l'originale e una copia della risorsa d'Archivio mediante un indirizzo e rappresenta così un link sempre attivo nel Catalogo della biblioteca digitale.

NBN (*National Bibliography Number*):

<<http://www.rinascimento-digitale.it/index.php?L=&ID=535&CERCA=NBN#535>>

È un Identificatore Persistente basato su URN per individuare in modo univoco una pubblicazione.

Il Progetto NBN Italia (NBN:IT) ha realizzato un servizio di Identificazione Persistente basato sullo standard NBN che assegna alle biblioteche nazionali la responsabilità esclusiva della gestione del sottodominio nazionale NBN:IT per la generazione degli Identificatori *NBN italiani* per:

- ✓ il deposito legale nazionale delle risorse digitali degli editori (art. 37 DPR 252/2006);
- ✓ il deposito delle Tesi di Dottorato in formato digitale presso le Biblioteche Nazionali Centrali di Firenze e Roma.

Per le risorse di *Proprietà Intellettuale* (PI) come articoli di riviste e set di dati scientifici

non basta di essere identificati solo da un *PID locale* che ne garantisca l'unicità e la persistenza a livello di *Repository*. Esse potrebbero essere meglio identificate attraverso l'utilizzo di un *Global Identifier* fornito di un robusto meccanismo di risoluzione che ne garantisca l'univocità e la persistenza della risorsa anche al di fuori del *Repository*.

□ **DOI (*Digital Object Identifier*) System:** <<http://www.doi.org/>>

Agenzie di registrazione di identificatori DOI:

- ✓ Crossref: <www.crossref.org>;
- ✓ mEDRA (*multilingual European DOI Registration Agency*): <www.medra.org>.

Il *Global Identifier* DOI:

- ✓ viene impiegato per l'identificazione duratura/persistente, all'interno di una rete digitale, di qualsiasi entità che sia oggetto di Proprietà Intellettuale (PI) e per associarvi i relativi dati di riferimento, i metadati, secondo uno schema strutturato ed estensibile;
- ✓ è un sistema internazionalmente condiviso per gestire riferimenti e citazioni di risorse correlate a *pubblicazioni scientifiche*.

DOI identifica la risorsa e non la sua localizzazione (URL). In caso di spostamento della URL, è sufficiente aggiornare la relativa informazione nella banca dati DOI per garantire la persistenza dei link costruiti anche in passato.

DOI è utilizzato per gestire:

- ✓ il “*Cross-linking*“ tra gli articoli di Riviste scientifiche online;
- ✓ il “*Forward-linking*“, un servizio che consente il rinvio alle pubblicazioni di un editore tramite le citazioni incorporate direttamente in una piattaforma (es. *Repository*);
- ✓ aggregazioni delle informazioni secondo aree tematiche omogenee, renderle disponibili da un unico punto di accesso (“*Multiple Resolution*“) identificato da un DOI

L'approccio del DOI è integrativo e non sostitutivo dei modelli già in essere. Ad esempio, a partire da un ISBN, è possibile creare – anche attraverso procedure automatiche – un link persistente DOI riferito alla risorsa identificata, rendendola così “azionabile in rete“ ovvero accessibile in Internet.

Fra i vari *Global Identifier* esistenti, il DOI è senza dubbio quello maggiormente affermato a livello internazionale e costituisce lo Standard più adottato nel mondo dei contenuti editoriali. I DOI possono essere abilitati per essere utilizzati in applicazioni di tipo *Linked Data*.

□ **Handle System:** <<http://www.handle.net/>>

Viene impiegato per assegnare e gestire Identificatori (*handles*) Persistenti agli oggetti digitali ed alle risorse presenti in rete.

“Sustainability of Digital FORMATS. Planning for Library of Congress Collections“:

<<http://www.digitalpreservation.gov/formats/intro/intro.shtml>>

Tra gli obiettivi di questo documento vi sono:

- ✓ supporto alla pianificazione strategica per quanto riguarda i Formati dei contenuti digitali con cui garantire la conservazione a lungo termine;
- ✓ identificare e descrivere i formati per la sostenibilità a lungo termine. Sviluppare strategie per sostenere questi formati incluso raccomandazioni riguardanti gli strumenti e la documentazione necessari per la loro gestione;
- ✓ identificare e descrivere formati non adatti per la sostenibilità a lungo termine, e sviluppare strategie per sostenere i contenuti incapsulati nei medesimi formati.

InterPARES (*International Research on Permanent Authentic Records in Electronic Systems*) Project: <<http://www.interpares.org/>>

È un Progetto istituito in Canada avente l'obiettivo di elaborare le conoscenze teoriche e metodologiche essenziali (Standard, Politiche, Strategie e Piani) sulla Conservazione a lungo termine di oggetti digitali.

NESTOR Project:

<http://www.dnb.de/DE/Wir/Kooperation/nesstor/nesstor_node.html>;

<http://www.langzeitarchivierung.de/Subsites/nesstor/DE/Home/home_node.html>

□ È una rete di cooperazione composta da 12 istituzioni dei paesi di lingua tedesca che promuovono la conoscenza sulla Conservazione del digitale.

□ “**NESTOR-Handbuch: Eine kleine Enzyklopädie der digitalen Langzeitarchivierung Cross-linking**“.(Manuale di NESTOR: Una piccola enciclopedia sul *cross-linking* della Conservazione a lungo termine):

<<http://www.nesstor.sub.uni-goettingen.de/handbuch/index.php>>:

Le tematiche del *Handbuch*:

- ✓ Politiche di Preservazione a lungo termine;
- ✓ Affidabilità dei contenuti e sicurezza dei sistemi;
- ✓ Modello OAIS;
- ✓ PREMIS (*PREservation Metadata: Implementation Strategies*) Standard;
- ✓ METS (*Metadata Encoding and Transmission Standard*);
- ✓ Formati per la Conservazione a lungo termine;
- ✓ Strategie della Conservazione digitale (migrazione, emulazione, microfilm);
- ✓ Accesso e reperimento. Identificatori persistenti;
- ✓ Hardware;
- ✓ Sistemi di *storage* per l'archiviazione a lungo termine;
- ✓ Aspetti legali;
- ✓ Procedure per l'archiviazione di diverse tipologie di oggetti (documenti di testo, immagini, multimedia, materiale didattico, documenti e dati di ricerca digitali)

V. anche **NESTOR: (Network of Expertise in Long-Term Storage of Digital Repositories)** (2009):

<<http://nbn-resolving.de/urn:nbn:de:0008-2010030806>>

CASPAR (Cultural, Artistic, and Scientific knowledge for PRESERVATION, Access and Retrieval) Project: <<http://www.casparpreserves.eu/>>

Il Progetto ha affrontato i temi su:

- ✓ un'ampia serie di aspetti relativi alla Conservazione di tutti i tipi delle informazioni digitali/digitalizzate e al loro futuro utilizzo;
- ✓ lo sviluppo degli strumenti sul modello OAIS per la conservazione, accesso e ricerca del sapere culturale, artistico e scientifico europeo negli ambienti digitali dinamici che richiedono soluzioni complesse per la preservazione a lungo termine;
- ✓ l'Agenda Digitale Europea (allegato “Progressi in materia di digitalizzazione e accessibilità *online* del materiale culturale e della conservazione digitale nell'UE”:

<<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:52008DC0513:IT:NOT>>

SHAMAN (Sustaining Heritage Access through Multivalent Archiving) Project:

<<http://shaman-ip.eu/>>

Il Progetto risponde alla necessità di Conservazione digitale definendo gli strumenti a supporto del Ciclo di vita dell'oggetto digitale, per mantenere le informazioni accessibili al di là dei limiti delle tecnologie, garantendo all'oggetto l'autenticità e l'integrità durante il suo ciclo di vita. La *Conservazione digitale* viene definita dal Progetto come la *comunicazione con il futuro*.

KEEP (Keeping Emulation Environments Portable) Project:

<<http://www.keep-project.eu/ezpub2/index.php>>

Molte copie di Software datati, dai primi videogiochi alle esclusive applicazioni Software professionali, sono state raccolte come reperti della prima era digitale.

Il Progetto si è occupato dello sviluppo dei servizi di emulazione per tradurre il codice e le istruzioni di un ambiente-macchina (qui obsoleto) in modo da farli girare correttamente in un altro.

Il Progetto ha sviluppato:

- ✓ una chiave ovvero un *emulatore* (un'applicazione Software, una macchina virtuale) con cui sbloccare il codice e vedere ciò che è basato sui Software obsoleti.

Questo emulatore funziona sui moderni computer, o persino sul Web, simulando l'Hardware e le componenti di sistema dei computer obsoleti;

- ✓ una serie di strumenti che aiuterà gli archivisti a estrarre dati da diversi tipi di "supporti" (il modo in cui il Software è stato "confezionato") e a convertire i dati in un formato di codifica comune e utilizzabile che sia indipendente dalla piattaforma;
- ✓ il *KEEP Media Transfer Tool Framework* (MTTF) che offre un metodo pratico per creare una "immagine" di un supporto Software e di memorizzarla su moderni supporti digitali in modo che possa essere usata da servizi di emulazione.

DPE (*Digital Preservation Europe*) Project: <<http://www.digitalpreservationeurope.eu>>

□ È un raggruppamento di numerosi sostenitori – per l'Italia aderisce anche l'AIB (Associazione Italiana Biblioteche) – per lo sviluppo delle Politiche comuni per la Conservazione del patrimonio digitale europeo.

□ DPE assieme con DCC (Digital Curation Center) ha sviluppato il *Toolkit DRAMBORA* (*Digital Repository Audit Method Based on Risk Assessment*):

<<http://www.repositoryaudit.eu/>>

Questo Toolkit è destinato a facilitare il controllo interno di *Repository* (per quanto riguarda la valutazione delle loro capacità relative alla Conservazione a lungo termine) l'identificazione di punti deboli, e il riconoscimento di punti di forza di *Repository* in questa prospettiva.

PLANETS (*Preservation and Long-term Access through Networked Services*) Project:

<<http://www.planets-project.eu/>>

È un progetto coordinato dalla *British Library* che raggruppa varie Biblioteche Nazionali, Università e soggetti privati per costruire servizi e strumenti pratici per garantire l'accesso a lungo termine ai Beni culturali e scientifici digitali.

PRONOM Project

<<http://www.nationalarchives.gov.uk/PRONOM/Default.aspx>>

□ È un progetto promosso dall'iniziativa inglese "*The National Archives*". Offre le informazioni tecniche sui formati dei file, sui Software e su altre componenti necessari a garantire l'accesso di lunga durata ai documenti elettronici e ad altre risorse digitali.

□ Nell'ambito del progetto è stato sviluppato un Software gratuito aggiornabile, che permette una facile identificazione dei formati dei file - DROID (*Digital Record Object Identification*), <<http://droid.sourceforge.net/>>

APARSEN (*Alliance Permanent Access to the Records of Science in Europe Network*)

European Project per "*excellence in Digital Preservation*": <<http://www.aparsen.eu>>

Partendo dai risultati di ricerca di altri progetti quali CASPAR, DPE e PLANETS, SHAMAN e PARSE.Insight, l'APARSEN integra le iniziative di ricerca attualmente esistenti in Europa sulla *Digital Preservation*, con la realizzazione dei significativi avanzamenti nei servizi personalizzabili per le risorse digitali scientifiche e culturali, migliorandone l'accesso e la fruibilità ai contenuti a lungo termine.

Progetto „Magazzini Digitali“ (Fondazione Rinascimento Digitale, Biblioteca Nazionale

Centrale di Firenze, Biblioteca Nazionale Centrale di Roma):

<<http://www.depositolegale.it/>>;

<<http://www.bncrm.librari.beniculturali.it/index.php?it/828/magazzini-digitali>>

□ Definizione della Conservazione come un servizio pubblico fornito da depositi digitali affidabili o fidati (*Trusted* o *Trustworthy Digital Repositories*) in grado di assicurare, per le risorse digitali depositate:

- ✓ la *Leggibilità* a livello di *bit* da parte di una macchina (*Viability*),
- ✓ la *Interpretabilità* a livello di formato (per esempio: pdf, doc ecc.) da parte di un elaboratore (*Renderability*),
- ✓ l'*Autenticità* (*Authenticity*) intesa come identità e integrità dell'oggetto digitale, e
- ✓ la effettiva *Disponibilità* (*Availability*) per le comunità designate (comunità di riferimento, interessate all'uso di quelle risorse).

□ Sviluppo di un Sistema per la Conservazione permanente dei documenti elettronici pubblicati in Italia e diffusi tramite rete informatica, in attuazione della normativa sul deposito legale (L. 106/2004, D.P.R. 252/2006).

CULTURAL HERITAGE on line – Trusted Digital Repositories.

(Conferenza internazionale 2012, Firenze):

<<http://www.rinascimento-digitale.it/conference2012-culturalheritageonline-materials.phtml>>

Gli argomenti principali:

- ✓ il caso italiano: legislazione e buone pratiche per la Conservazione del digitale e le soluzioni principali per la costruzione di *Trusted Repository*;
- ✓ il futuro per la collaborazione internazionale sul tema della Conservazione del digitale;
- ✓ lo Standard OAIS e la certificazione di *Trusted Digital Repositories*.

Bibliografia selettiva in relazione al tema 'Digital Preservation'

□ **Digital Curation Bibliography: "Preservation and Stewardship of Scholarly Works":** <<http://digital-scholarship.org/dcpb/dcb.htm>>. Il supplemento con la nuova bibliografia (2013): <<http://digital-scholarship.org/dcpb/dcpb.htm>>

□ **ARIADNE (Web Magazine for Information Professionals):**

<<http://www.ariadne.ac.uk/taxonomy/term/1353>>

□ **"Long-term preservation of Enhanced Publications" (DRIVER Technology Watch):**

<http://www.driver-repository.eu/component/option,com_jdownloads/Itemid,83/task,viewcategory/catid,8/>

□ **Wikipedia (l'enciclopedia libera):**

<http://it.wikipedia.org/wiki/Preservazione_digitale>

4.6. Digitalizzazione

Per digitalizzazione s'intende il processo di trasformazione dello stato originale di oggetti, quali documenti cartacei, di pellicole fotografiche, cinematografiche, microfiche o microfilm, lastre o pellicole radiologiche, cassette e nastri magnetici audio o video, dal formato analogico a quello digitale, interpretabile da computer. Il formato digitale è un codice binario in cui tutto è rappresentato da combinazioni di zero od uno, quindi da stati del tipo acceso/spento.

Un Archivio Istituzionale (*Institutional Repository*, IR) – oltre a rappresentare un luogo virtuale per l'auto-archiviazione dei contenuti di ricerca prodotti direttamente in formato digitale - può includere nelle loro collezioni anche i contenuti digitalizzati¹⁰:

¹⁰ La digitalizzazione dei materiali protetti da *COPYRIGHT* è possibile invocando l'Istituto del *fair use* e rendendo accessibili all'utenza solo brevi frammenti di testo, detti *snippet*. L'idea è quella che gli utenti potessero ricercare

*A Repository [...] can also contain the Institutions Digital Library, the collection of **printed and manuscript documents**, public archives, and graphic material, originating from the Institution or elsewhere, that the University **has converted to digital form** for use within the University, and generally available to anyone [...] The main reasons to **digitize** are to **enhance access and improve preservation**. It is a means of creating a database of materials like Theses and Dissertations and other resources that worth preservation [Eke, 2011, online, non impaginato].*

Un esempio dell'IR che ospita anche i prodotti digitalizzati è "Labordoc" (implementato sul Software *Open Source Invenio*). Esso contiene più di 80.000 risorse in versione digitale dell'*International Labor Organization* a partire dal 1919 [V. sito: Labordoc].

"Ma se non viene *digitalizzato* e reso *accessibile online*, questo patrimonio domani potrebbe non riuscire a svolgere il suo *ruolo determinante* nella *futura geografia della conoscenza*" [De Robbio, 2007, p.284].

L'espressione italiana:

- ✓ "*Digitalizzazione*" indica tendenzialmente il processo di "*Conservazione digitale*" di documenti analogici, mentre
- ✓ "*Conservazione del documento*" (della memoria/ informazione/ risorsa *digitale*) è la più comune traduzione dell'espressione *Digital Preservation* [Faiella, 2000, p.64]

La digitalizzazione insieme con conservazione del materiale negli Archivi Istituzionali non deve essere

la singola soluzione del COME, piuttosto che del CHI o del CHE COSA, a fornire una rete di sicurezza in grado di garantire l'*ACCESSO alla documentazione elettronica nel FUTURO*, che si potrà costruire invece solo assicurando STRATEGIE EFFICACI ed INFRASTRUTTURE organizzative ADEGUATE a fronteggiare i problemi attuali e quelli futuri [Faiella, 2000, p.68]

I dati del file digitalizzato (scansionato) possono essere "catturati" tramite:

- ✓ *BarCode* (1D e 2D);
- ✓ *Field OCR (Optical Character Recognition)* per il riconoscimento di parole-numeri su zone specifiche;
- ✓ *Free Form OCR* (riconoscimento "semantico" del campo: ricerca nel documento del campo voluto e cattura del contenuto);
- ✓ *OMR (Optical Mark Recognition)* per il riconoscimento segni o *box* anneriti;
- ✓ *ICR* per il riconoscimento della scrittura umana
[V. sito: "Archiviazione documentale: 2C Solution"].

I file digitalizzati possono essere:

a) ulteriormente *elaborati* per eventuali attività di:

- ✓ pulizia e riconoscimento del testo (OCR);
- ✓ riconoscimento delle marcature su questionari, *check-list* e simili tramite OMR;
- ✓ creazione o conversione file in formato Acrobat (PDF e PDF/A);
- ✓ masterizzazione di file immagine per una successiva elaborazione CAD (*Computing Aided Design*)

b) *indicizzati* in modo:

- ✓ manuale da professionisti del settore o
- ✓ semi-automatico o automatico.

Per l'indicizzazione automatica basata su parole-chiave del materiale digitalizzato, negli IR in Italia può essere impiegato il Thesaurus del Nuovo Soggettario o un Thesaurus Specializzato creato a livello locale e basato sulle voci del Thesaurus del Nuovo Soggettario e con esso interoperabile. Tale prospettiva di indicizzazione dei documenti

liberamente il materiale e visualizzare i risultati in un gruppo (*cluster*) di 2-3 frasi in cui compare l'elemento ricercato entro un testo connesso con il mercato editoriale (per l'eventuale acquisto qualora in commercio) o presente in qualche Biblioteca (se fuori dal circuito commerciale. In questo modo si dà accesso a libri che altrimenti non si sarebbero mai potuti ottenere – come è in larga parte tutt'ora. A consultazione per intero, previa digitalizzazione, dei libri di *pubblico dominio* (70 anni dopo la morte dell'autore) non presenta aspetti particolarmente problematici.

potrebbe rappresentare la migliore soluzione per *uniformare il trattamento semantico* delle risorse documentarie istituzionali all'interno di Enti diversi, come pure per *correlare fra loro IR e OPAC, strutture documentarie* sicuramente raccordabili e che possono *superare ogni forma di incomunicabilità* [Guerrini, 2010a, p.48].

In seguito saranno esposte informazioni, elencate fonti, e forniti rimandi di approfondimento relativamente al tema del presente paragrafo.

Supporto al processo OCR (*Optical Character Recognition*)

“Sperimentazione e analisi di tecniche di Riconoscimento dei Caratteri (OCR) per la digitalizzazione di documenti” [Caruso]:

https://www.cnipa.gov.it/site/files/Refo2_RS_11.0_Caruso.pdf

La relazione viene messa a disposizione degli esperti e specialisti del settore, nel mondo accademico ed in quello della ricerca industriale.

Software *FineReader* per Linux: http://www.abbyy.com/ocr_sdk_linux/

Tale Software puo' essere integrato nel sistema per realizzare un servizio di OCR scalabile (es. durante l'*ingestion*, la *dematerializzazione*). Le caratteristiche del *FineReader*:

trascrizione del testo, indicizzazione, ricerca *full-text*;

capacità di trascrizione testo anche se organizzato in colonne o in modo strutturato nella pagina e diversi formati di *output*;

trascrizione automatica delle immagini (TIFF, JPG) in testo (file di testo, XML, PDF con testo selezionabile) e ricerca *full-text* tramite trascrizione XML;

riconoscimento multi-lingua tramite l'utilizzo di dizionari a supporto di più di 190 lingue nazionali antiche e moderne.

Il testo ottenuto con OCR è una base essenziale per ulteriori analisi semantiche che permettono di identificare concetti all'interno dei testi.

Le “Linee Guida sulla digitalizzazione”.

Il Gruppo di lavoro dell'Archivio digitale PHAIDRA, Padova:

<https://phaidra.cab.unipd.it/static/linee-guida-digitalizzazione.pdf>

Le Linee descrivono le procedure coinvolte nella digitalizzazione “*in-house*” o “*in outsourcing*”, in coerenza con le Buone Pratiche e Standard nazionali e internazionali per una riproduzione qualitativa di documenti.

In linea di principio, tutti i Progetti di digitalizzazione devono essere coerenti con uno o più dei seguenti scopi generali:

- ✓ valorizzare il patrimonio documentale rilevante a fini scientifici e culturali;
- ✓ ampliare l'accesso del pubblico al medesimo patrimonio documentale;
- ✓ favorire gli studi interdisciplinari sul materiale digitalizzato e promuovere la collaborazione tra diverse istituzioni del territorio;
- ✓ promuovere la conoscenza di collezioni locali o uniche, mediante la loro diffusione su larga scala;
- ✓ creare collezioni virtuali attraverso l'integrazione di vari formati o di materiali distribuiti in luoghi diversi;
- ✓ limitare la consultazione diretta di documenti originali in particolari condizioni di criticità;
- ✓ facilitare l'accesso a materiale di difficile fruizione;
- ✓ assicurare che il materiale documentale sia disponibile alle future generazioni di lettori;
- ✓ velocizzare la gestione documentale di un'Istituzione.

Il processo di digitalizzazione di documenti cartacei permette di ridurre molto lo spazio dove

immagazzinarli e, inoltre, aumenta di molto la loro vita.

Il passaggio dei documenti al digitale permette una loro migliore Classificazione e di conseguenza una semplice e veloce reperibilità.

“Guidelines for digitization projects” (International Federation of Library Associations and Institutions/IFLA): <<http://archive.ifla.org/VII/s19/pubs/digit-guide.pdf>>

- ✓ Criteri per la selezione dei documenti per la digitalizzazione;
- ✓ Conversione dei documenti in digitale;
- ✓ Controllo di qualità;
- ✓ Gestione delle collezioni dei documenti digitalizzati;
- ✓ Aspetti legali.

In certi stati, come la Francia, il *Ministero della ricerca* si fa carico di tutelare le Biblioteche digitali nel caso gli autori reclamassero diritti per il materiale digitalizzato, in quanto anche la Francia, essendo come l'ITALIA in regime di DIRITTO d'AUTORE e non di Copyright, necessiterebbe anche dell'autorizzazione dell'autore prima di mettere a disposizione il lavoro nella versione digitale [De Robbio, 2007, p.132];

- ✓ Costi;
- ✓ Programmazione di risorse umane;
- ✓ Sviluppo e mantenimento di interfacce web per documenti digitalizzati;
- ✓ Preservazione dei contenuti digitali.

State-of-the-art Tools for Text Digitisation:

<<http://tpdl2013.upatras.gr/tut-digit.php>>

- ✓ Tecniche di digitalizzazione del testo;
- ✓ Identificazione dei punti di forza e di debolezza di usabilità di strumenti di digitalizzazione esistenti;
- ✓ Raggiungimento di una migliore conoscenza sui requisiti e gli effetti dell'integrazione degli strumenti di digitalizzazione nel flusso di lavoro di produzione dei documenti digitali.

Linee Guida e tecniche necessari al coordinamento dei Progetti di Digitalizzazione.

Istituto centrale per il Catalogo Unico delle Biblioteche italiane e per le informazioni bibliografiche (ICCU):

<http://www.iccu.sbn.it/opencms/opencms/it/main/bdi/linee_guida_standard/>

Tali Linee di indirizzo sono necessarie agli Enti che stanno già programmando, eseguendo o gestendo Progetti di digitalizzazione e funzionali alla creazione di un sistema di servizi informativi e di valorizzazione del patrimonio culturale in Italia.

□ “Linee di indirizzo per i progetti di digitalizzazione del Materiale Fotografico”:

<http://www.iccu.sbn.it/upload/documenti/Linee_guida_fotografie.pdf>

□ “Linee Guida per la digitalizzazione del materiale Cartografico”:

<http://www.iccu.sbn.it/opencms/export/sites/iccu/documenti/linee_guida_digit_cartografia_05_2006.pdf>

□ Linee Guida per la digitalizzazione di bandi, manifesti e fogli volanti:

<http://www.iccu.sbn.it/opencms/export/sites/iccu/documenti/linee_guida_bandi_sett.2006.pdf>

□ Linee Guida tecniche per i programmi di creazione di contenuti culturali digitali:

<http://www.minervaeurope.org/publications/Linee_%20guida_%20tecniche.pdf>;

<http://www.minervaeurope.org/publications/technicalguidelines_it.htm>

□ “Tecniche di Digitalizzazione - Istituto Centrale Beni Sonori e Audiovisivi”:

<<http://ceaih.dds.it/index.php?it/137/tecniche-di-digitalizzazione-istituto-centrale-beni-sonori-e-audiovisivi>>

“Ministerial network for valorizing activities in digitization.

Progetto MINERVA: <<http://www.minervaeurope.org/publications/buonepratiche.htm>>

Identificazione delle buone pratiche e dei centri di competenza per la digitalizzazione.

“Collection Digitization Policy”. National Library of Australia:

<<http://www.nla.gov.au/policy/digitisation.html>>

Questa Politica viene monitorata annualmente dalla “Digitisation & Photography Branche“ e dal “Library’s Collection Development and Management Committee“ per verificare i cambiamenti nelle politiche delle Biblioteche relative alla pianificazione strategica che tiene conto degli sviluppi del settore in attività di digitalizzazione.

“The NINCH Guide to Good Practice in the Digital Representation and Management of Cultural Heritage Materials”: <<http://www.nyu.edu/its/humanities/ninchguide/>>

Buone pratiche nella digitalizzazione e nella gestione dei materiali del patrimonio culturale:

- ✓ Project Planning;
- ✓ Selezione dei materiali: un processo iterativo;
- ✓ Gestione dei diritti;
- ✓ Digitalizzazione e codifica dei testi;
- ✓ Cattura e gestione delle immagini;
- ✓ Cattura di audio/video e loro gestione;
- ✓ Controllo e garanzia di qualità’;
- ✓ Distribuzione;
- ✓ Sostenibilità’: modelli per il finanziamento a lungo termine;
- ✓ Valutazione di Progetti dagli utenti;
- ✓ Gestione di risorse digitali;
- ✓ Preservazione di risorse digitali.

“Manuale virtuale per la Progettazione digitale” .

Regione Toscana. Settore biblioteche, archivi, istituzioni culturali e catalogo dei beni culturali:

<http://www.cultura.toscana.it/biblioteche/biblioteca_digitale/manuale/>

- ✓ Definizione del Progetto;
- ✓ Produzione;
- ✓ Metadati;
- ✓ Pubblicazione sul web;
- ✓ Conservazione;
- ✓ Costi.

“Minimum Digitization. Capture Recommendations”.

The Association for Library Collections and Technical Services Preservation and

Reformatting Section: <http://connect.ala.org/files/43293/MCTF_Draft_Recommendation.pdf>

Le Raccomandazioni per le Biblioteche che hanno intenzione di digitalizzare contenuti disponibili (“static media” e “time-based media”) e conservarli a lungo termine. Le Raccomandazioni hanno l’obiettivo di dare indicazioni di base in modo che il lavoro di digitalizzazione possa essere fruibile anche in futuro.

“Il patrimonio Culturale Europeo: basta un clic.

Progressi in materia di digitalizzazione e accessibilità on-line del materiale culturale e della conservazione digitale nell’UE”.

Comunicazione della Commissione al Consiglio, al Parlamento europeo:

<<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:52008DC0513:IT:NOT>>

La presente comunicazione illustra i progressi realizzati fino a questo momento e le misure da adottare per la creazione del Portale EUROPEANA: una biblioteca digitale europea che offre l’accesso a oltre 5 milioni di libri, giornali, filmati, carte, fotografie e documenti digitalizzati, provenienti da oltre 1000 istituzioni culturali: <<http://www.europeana.eu/>>

Comunicazioni della Commissione Europea sulla digitalizzazione

□ **“Raccomandazione sulla *“Digitalizzazione e sulla conservazione digitale nel lungo periodo”*”:**

<<http://www.libereurope.eu/news/european-commission%E2%80%99s-recommendation-on-digitalisation-and-digital-preservation>>

La Commissione Europea invita gli stati membri di concentrare le risorse per la digitalizzazione del materiale culturale rendendolo disponibile attraverso EUROPEANA.

□ **Raccomandazione sulla *“Digitalizzazione e l’accessibilità in rete dei materiali culturali e sulla conservazione digitale”*”:**

<<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2011:283:0039:0045:IT:PDF>>

La Commissione sottolinea la necessità di rendere accessibili i materiali digitalizzati attraverso il Portale EUROPEANA, di garantire l’uso di norme comuni per la digitalizzazione, per favorire l’interoperabilità, di rendere disponibili a titolo gratuito i metadati esistenti, con l’obiettivo finale di rendere accessibili in EUROPEANA tutte le principali opere in pubblico dominio pubblicate in Europa, a partire dal 2015 [Casella, 2013, pp. 227, 228].

“Digitizing Dissertations for an Institutional Repository: A Process and Cost Analysis”:

<<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2479051/>>

La descrizione di un Progetto di digitalizzazione delle Dissertazioni per un Archivio Istituzionale. *“Each Dissertation was:*

- ✓ *scanned,*
- ✓ *reviewed for Quality Control,*
- ✓ *enhanced with a Table of Contents,*
- ✓ *processed through an Optical Character Recognition function, and*
- ✓ *added to the Institutional Repository”.*

Alcune semplici Guide online a supporto dei processi di digitalizzazione di testi e immagini

□ **“Uso dello scanner per digitalizzare testi ed immagini”:**

<<http://www.multimedididattica.it/dm/scanner/scanner.htm>>

□ **“Come digitalizzare un testo con lo scanner e convertirlo in .doc, .txt, .pdf, ecc.”:**

<<http://www.programmifree.com/guide/digitalizzare-scanner.htm>>

Alcuni casi d’uso della rappresentazione dei contenuti digitalizzati

□ **“Integrating Digitized material into an Institutional Repository: the case of “SOMNI” and “EUROPEANA REGIA” at the Universitat De València”:**

<<http://indicotest.eblida.org/getFile.py/access?contribId=57&sessionId=12&resId=0&materialId=0&confId=22>>

La descrizione di un Progetto di digitalizzazione dei Manoscritti e il loro deposito nell’Archivio Istituzionale dell’Università di València (Figura 42) .

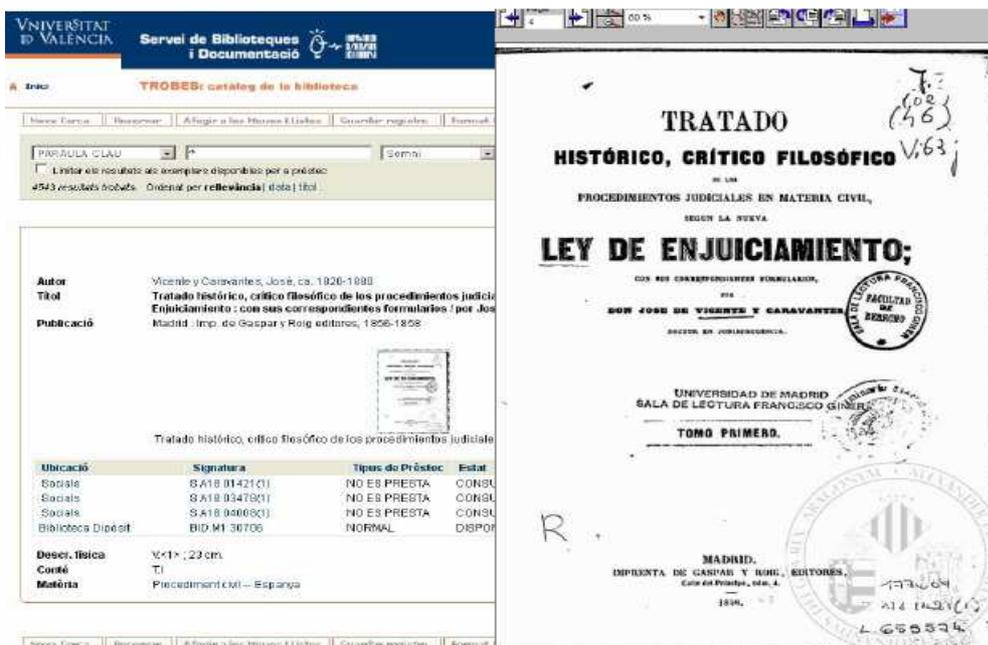


Figura 42. Rappresentazione dei digitalizzati nell'Archivio Istituzionale dell'Università Valentia

□ L'Archivio Istituzionale *DSpace@Cambridge*:

<<http://www.lib.cam.ac.uk/repository/theses/>>

prevede la possibilità del deposito delle Tesi disponibili originalmente nelle copie cartacee ("Deposit of old Thesis") (Figura 43), dopo essere digitalizzate e convertite in determinati formati.

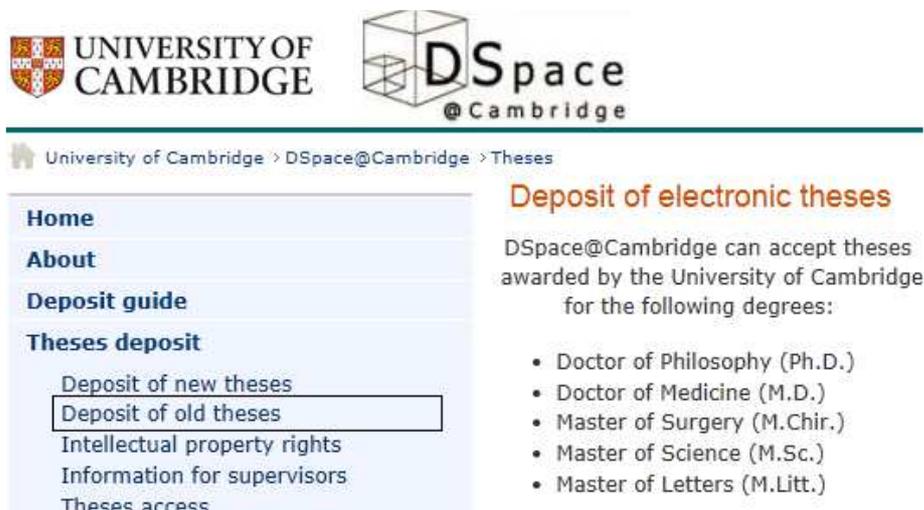


Figura 43. "Deposit of old Thesis" nell'Archivio Istituzionale *DSpace@Cambridge*

□ L'Archivio Istituzionale *Phaidra* (Università di Vienna):

<<https://fedora.phaidra.univie.ac.at/fedora/get/o:69507/bdef:Book/view>>

La rappresentazione e la consultazione dei contenuti di una Tesi digitalizzata (Figura 44)

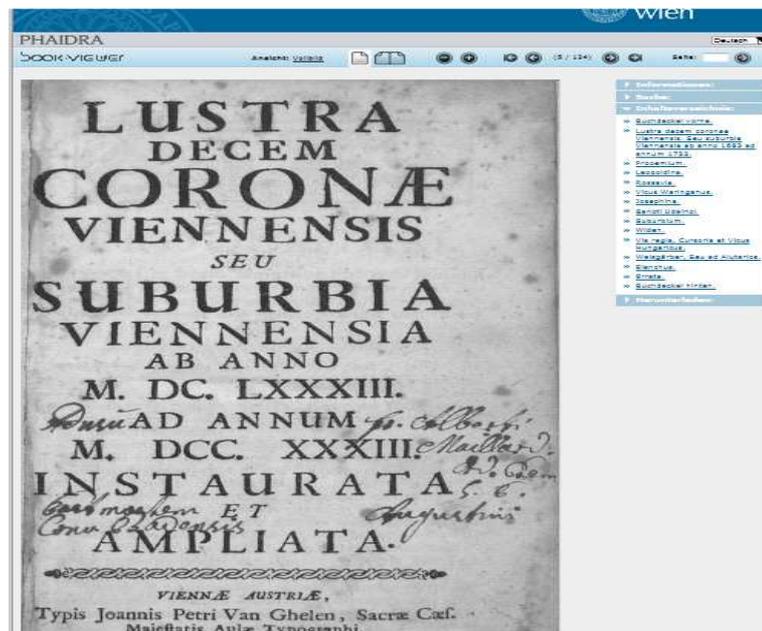


Figura 44. Rappresentazione dei contenuti di una Tesi digitalizzata nell'Archivio PHAIDRA (Università di Vienna)

4.7. Scoperta delle risorse: paradigmi di ricerca e navigazione

l'Archivio dovrebbe divenire lo *spazio di lavoro quotidiano* per i ricercatori e per gli amministratori della ricerca. Per far questo bisognerebbe dotare l'Archivio di un'interfaccia amichevole che contenga tutte le informazioni necessarie in una struttura semplice e agevole, con un agile sistema di supporto per gli utenti /autori e gli utenti / fruitori (*Help Desk/ Faq/ Help Online* ecc.)

[V. sito: "Linee Guida per gli Archivi Istituzionali", p.17]

Le procedure relative alla SCOPERTA delle risorse sono:

- ✓ Cercare un'entità che corrisponda a criteri stabiliti usando un attributo o una relazione dell'entità come criterio di ricerca;
- ✓ Identificare un'entità, cioè confermare che l'entità rappresentata corrisponda all'entità cercata, o distinguere tra due o più entità con caratteristiche simili;
- ✓ Controllare la forma dell'intestazione per assicurarsi che sia usata coerentemente in un sistema informativo al fine di supportare la funzione di raggruppamento;
- ✓ Collegare ovvero stabilire o chiarire la relazione tra un'entità e l'altra cioè stabilire la relazione tra due o più intestazioni autorizzate (collegamento "Vedi Anche") o tra intestazioni varianti e intestazione autorizzata (collegamento "Vedi") specificandone, se necessario, la natura

[Weston, Agnese, 2003, pp.6-7]

Tra gli strumenti a supporto della strutturazione della conoscenza e ricerca dei contenuti correlati:

- ✓ Vocabolari controllati che aiutano in precisione e accuratezza della ricerca;
- ✓ "Ontology-based Image Retrieval": dal recupero di base al recupero delle parti delle immagini tramite la loro segmentazione mediante applicazione di Ontologie, tecniche di apprendimento automatico e metodi statistici;
- ✓ LINKED DATA

DATA and traditional AUTHORED WORKS are starting to INTEGRATE and RELATE to each other in complex ways. There is an increasing sense of contributing to the COLLECTIVE KNOWLEDGE BASE, to the corpus of knowledge [through]

LINKED DATA [...] Scientists can now REPRODUCE an experiment with the data or REANALYSE them in DIFFERENT WAYS, whereas previously the data collected in the course of research might never really be used again. This idea of REUSE is key here; reuse ADVANCING SCIENTIFIC DISCOVERY [Stevenson, 2009, online, non impaginato]

In seguito saranno esposte informazioni, elencate fonti, e forniti rimandi di approfondimento relativamente al tema del presente paragrafo.

“Improving the *QUALITY* of metadata in Eprint Archives”:

<<http://www.ariadne.ac.uk/issue38/guy/>>

I *metadati* (indici descrittivi) di alta qualità devono tener conto dei *requisiti funzionali* (l' idoneità allo scopo) del Sistema al quale sono di supporto, tra cui sono i *requisiti funzionali* che riguardano RICERCA (Semplice e Avanzata) e NAVIGAZIONE.

Gli utenti dell' Archivio Istituzionale devono essere in grado di:

1) ricercare il materiale depositato per:

- Titolo
- Nome dell' Autore
- Parole chiave
- Tipo di documento
- Dati della pubblicazione
- Titolo del convegno
- Anno

2) Scorrere o sfogliare (Browse) i contenuti per:

- Anno
- Struttura universitaria
- Ultimi documenti depositati

3) Consultare i contenuti collegati fra loro in relazione:

- allo stesso Convegno
- alla stessa Pubblicazione

4) Filtrare i contenuti per:

- Anno
- Tipo di documento

il gestore del sistema dovrà avere cura che la *Lista organizzata* risponda nella maniera più semplice e *direttamente comprensibile alle necessità di ricerca dell'utente*: dovrà prestare quindi attenzione alla strutturazione di un' *interfaccia* in modo che essa permetta la *massima facilità d'uso* e utilizzi le riflessioni elaborate negli ultimi anni sul tema degli OPAC [*On-line Public Access Catalogue*] (per esempio, Web 2.0) e, prima ancora, dei *Principi di Catalogazione internazionali* dell'IFLA (ICP)

che si riferiscono in primo luogo a *Registrazioni bibliografiche di Autorità* e a cataloghi correnti di Biblioteca, ma che possono anche applicarsi a Bibliografie e *Archivi di dati* creati da biblioteche, Archivi, musei e altre istituzioni, compresi gli *Archivi Istituzionali* [Guerrini, 2010 a, pp. 55-56]

[V. sito Dichiarazione di Principi Internazionali di Catalogazione]

La *manutenzione* della *Lista organizzata* e l' *aggiornamento* costante del *Software* rappresentano impegni *irrinunciabili* ai quali le Istituzioni e i responsabili non possono sottrarsi [...]

L'OPAC e l'ARCHIVIO ISTITUZIONALE, inoltre, dovrebbero risultare *interoperabili* tramite il Portale di Ateneo. Essi rimangono due strumenti distinti, che contengono *dati* particolari e *garantiscono la ricerca*, ma possono essere efficacemente inseriti in uno *spazio condiviso* (strumenti di *META-SEARCH*) all'interno del Portale di Ateneo

Importanza di METADATI SEMANTICI e *BROWSING* per SOGGETTO

□ Tra i servizi che possono contribuire a ELEVARE i livelli di QUALITÀ degli IR rientrano senza dubbio l'assegnazione di METADATI SEMANTICI ai contributi depositati e il potenziamento delle funzioni del *BROWSING* per SOGGETTO.

Si dovrebbero utilizzare DESCRITTORI che, riferiti al SOGGETTO dei documenti cioè al loro "tema di base", ne permettano il reperimento. In Italia, attualmente, l'offerta di accessi semantici per questo tipo di materiale è molto diversificata.

Alcuni IR consentono l'inserimento e pertanto la ricerca solo tramite la declaratoria dei SETTORI SCIENTIFICI CO-DISCIPLINARI stabiliti dal MiUR, mentre altri IR stanno sperimentando l'impiego di SOGGETTARI elaborati e adottati da agenzie catalografiche e pertanto redatti secondo un linguaggio documentario consolidato, oppure impiegano Schemi di CLASSIFICAZIONE di varia natura. L'IR del l'Università di Trento, per esempio, ha adottato le LCSH, *Library of Congress Subject Headings*

[Guerrini, 2010 a, p. 46]

□ L'Archivio aperto disciplinare "E-LIS", ad esempio, ha predisposto uno Schema di Classificazione JITA.

□ Dalla realtà pratica di VOCABOLARI CONTROLLATI e Schemi di CLASSIFICAZIONE che rendono improponibile qualunque forma di omologazione "scaturisce lo stimolo a VALORIZZARE le ANALOGIE e a costruire su di esse dispositivi che permettano APPROCCI TRASVERSALI agli strumenti di RICERCA creati da ciascuno"[Weston, Galeffi, 2003, p. 8].

□ [V. la sezione del *ToolKit* "Associazione di metadati con valori semantici"]

La Ricerca a FACCETTE: Aspetti teorici e pratici [Marcato, 2010]:

<http://tesi.cab.unipd.it/26030/1/La_Ricerca_a_Faccette.pdf>

□ La Ricerca a FACCETTE e' una *tecnica per l'accesso* alle informazioni le cui origini teoriche risalgono ai primi decenni del Novecento. Tuttavia la sua applicazione in ambito informatico è stata concretizzata solo recentemente.

□ *Ranganathan* introdusse nel 1933 un sistema di Classificazione che rispecchiava la DECOMPOSIZIONE di un CONCETTO in FACCETTE (*isolates*), in cinque macro categorie mutualmente esclusive: PERSONALITÀ, MATERIA, ENERGIA, SPAZIO, TEMPO (PMEST).

□ Tecniche di *INFORMATION RETRIEVAL* incentrate sulla CLASSICAZIONE a FACCETTE (o multidimensionale):

✓ RICERCA PARAMETRICA.

L'interfaccia di ricerca permette agli utenti di *formulare interrogazioni* specificando le varie condizioni sui *valori* delle *faccette*. La *Query* risultante sarà composta dagli operatori booleani AND e OR;

✓ NAVIGAZIONE A FACCETTE.

Permette all'utente di sviluppare *l'interrogazione progressivamente*, vedendo di volta in volta *l'effetto* che la *scelta di un valore su una faccetta* produce sui *valori disponibili* per le *altre faccette*.

✓ RICERCA A FACCETTE

si basa sull'esistenza di un *qualche insieme di documenti* organizzato secondo un *sistema di Classificazione multidimensionale*.

□ Principi per la SCELTA delle FACCETTE:

Differenziazione;

✓ Rilevanza;

✓ Accertabilità;

✓ Permanenza;

✓ Omogeneità;

✓ Mutua esclusività;

✓ Categorie Fondamentali

(Figura 45).

□ Ogni FACCETTA poi è strutturata GERARCHICAMENTE per divisione e le varie GERARCHIE possono essere COMBinate tra loro per formare ARGOMENTI COMPOSITI multidimensionali.

□ Se la parte strutturata di un DOCUMENTO è stata progettata per essere conforme a un sistema di CLASSIFICAZIONE a FACCETTE, allora è possibile COMBINARE:

✓ la RICERCA TESTUALE, applicata alla parte non strutturata del documento, con

✓ la navigazione a FACCETTE sui METADATI.

La disponibilità di Metadati strutturati a faccette è un requisito irrinunciabile.



Figura 45. Faccette a supporto delle Categorie Fondamentali per la ricerca e navigazione nel Portale DRIVER

Software *AquaBrowser Library*.

Ricerca e navigazione per associazioni semantiche:

<<http://win.cenfor.net/products/database/bowker/AquaBrowserBrochure.pdf>>

<http://www.cenfor.net/index.php?option=com_content&view=article&id=87&Itemid=6>

□ *AquaBrowser Library* nasce da uno studio approfondito condotto sulla PSICOLOGIA del RICERCATORE di fronte alle TECNICHE (o ABITUDINI) di RICERCA su fonti sempre più vaste ed eterogenee.

□ *AquaBrowser Library* è un potente motore di ricerca ed è un visualizzatore che consente il recupero di dati su fonti differenti (Catalogo Bibliografico, Banche Dati, Risorse Elettroniche, Documenti Digitali) in una modalità semplice, intelligente ed assolutamente innovativa (“SEARCH” – “DISCOVER” – “REFINE”).

□ La ricerca e navigazione sono basate su una rete di ASSOCIAZIONI SEMANTICHE (associazioni tra termini, varianti ortografiche, termini provenienti da Vocabolari Controllati, Clasificazioni, traduzioni in altre lingue) PRODOTTE nella fase di CONFIGURAZIONE ed INDICIZZAZIONE dei dati contenuti nel database del sistema.

□ L'APPROCCIO più diffuso tra chi opera ricerche in ambiti vasti ed eterogenei è quello che viene ormai definito “GOOGLE-LIKE” intendendo con tale espressione l'ATTITUDINE a RICERCHE SEMPLICI, su fonti non necessariamente “strutturate” e senza una necessaria esperienza e competenza tecnologica nell'uso di risorse differenti.

Esempio di implementazione: sito del Museo Galileo (Figura 46):

<<http://www.museogalileo.it/>>



Figura 46. Ricerca e navigazione per associazioni semantiche tramite Software *AquaBrowser Library*. Sito del *Museo Galileo*

[V. la sezione del *ToolKit* “Panorama delle piattaforme”].

“*User interface models for NEW PARADIGMS to SCHOLARLY SEARCH and BROWSING*”. Progetto VOA3R:

<<http://agroknow.redpanda.gr/content/d33uahdraft20110221-user-interface-models-new-paradigms-scholarly-search-and-browsing>>

<<http://win.cenfor.net/products/database/bowker/AquaBrowserBrochure.pdf>

<http://www.cenfor.net/index.php?option=com_content&view=article&id=87&Itemid=6>

L'ANALISI dei PARADIGMI di RICERCA e NAVIGAZIONE esistenti e la presentazione di un'architettura di informazione e di modelli di interfaccia per la ricerca e navigazione di avanguardia.

Esempi delle funzionalità di Navigazione e Ricerca Semantica

□ **ONTOLOGY NAVIGATION:**

<http://www.organic-edunet.eu/index.php?option=com_navigational&view=navigational&Itemid=89>

La Navigazione Semantica *ONTOLOGY NAVIGATION* (Figura 42) avviene attraverso le INTERFACCE GRAFICHE sulla base di ONTOLOGIE e/o TASSONOMIE implementate attraverso i *TOPICS* definiti tramite THESAURI collegati ai *dataset* che definiscono le PERSONE e altre ENTITÀ.

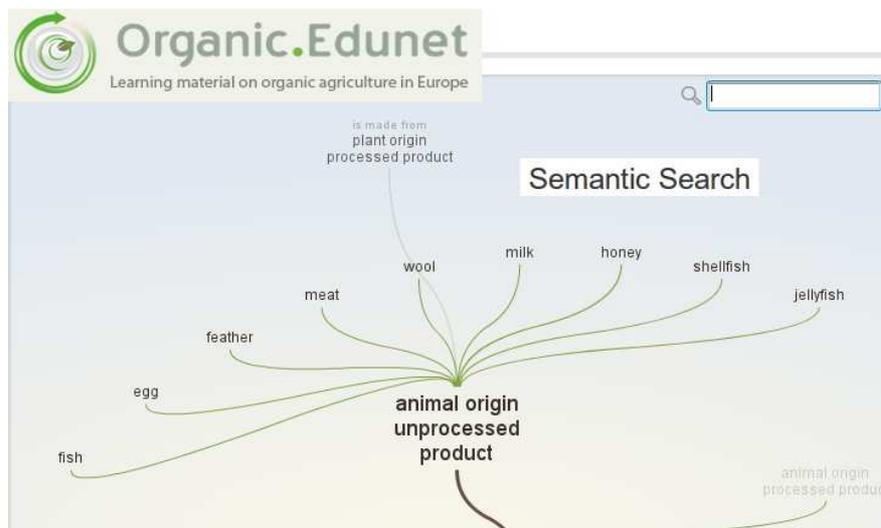


Figura 47. Interfaccia di navigazione semantica “*ONTOLOGY NAVIGATION*”

□ *ATLAS of SCIENCE* [V. sito: “*Scimago Reserach Group*”]

Questo tipo di Navigazione Semantica (Figura 48) si basa sulle informazioni estratte dalle pubblicazioni e correlate in base all’affinità’ semantica.

Le informazioni correlate possono riguardare:

- ✓ persone o gruppi di ricerca;
 - ✓ persone o gruppi di ricerca e loro pubblicazioni in una determinata area;
 - ✓ dati di ricerca in un determinato campo e in campi trasversali;
 - ✓ progetti europei tra paesi europei in base ai progetti comuni;
- campi del sapere nella/e scienza/e.

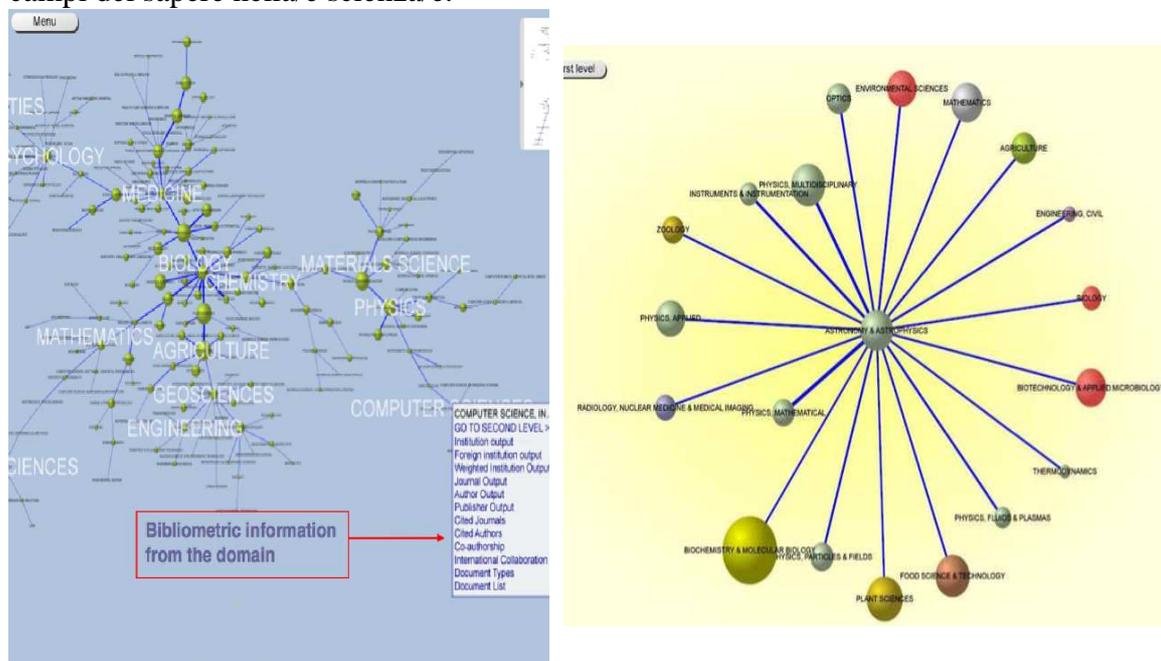


Figura 48. *ATLAS of SCIENCE*: navigazione semantica per campi del sapere nelle scienze

□ **Navigazione per *TIMELINE*** (Figura 49) potrebbe essere particolarmente utile per rappresentare:

- ✓ l'andamento di studio di una scienza, trasmettendo un senso di cambiamento nel corso del tempo;
- ✓ pubblicazioni e temi di ricerca in ordine di data;
- ✓ autori più citati per temi di ricerca;
- ✓ lavori di ricerca in corso;
- ✓ tappe dell'evoluzione di un articolo;
- ✓ progressi dei progetti;
- ✓ stato di sviluppo di un esperimento e così via.

I meccanismi di "Filtro" possono essere definiti a seconda delle esigenze di ricerca [V. sito: "*Timeline Template*"]

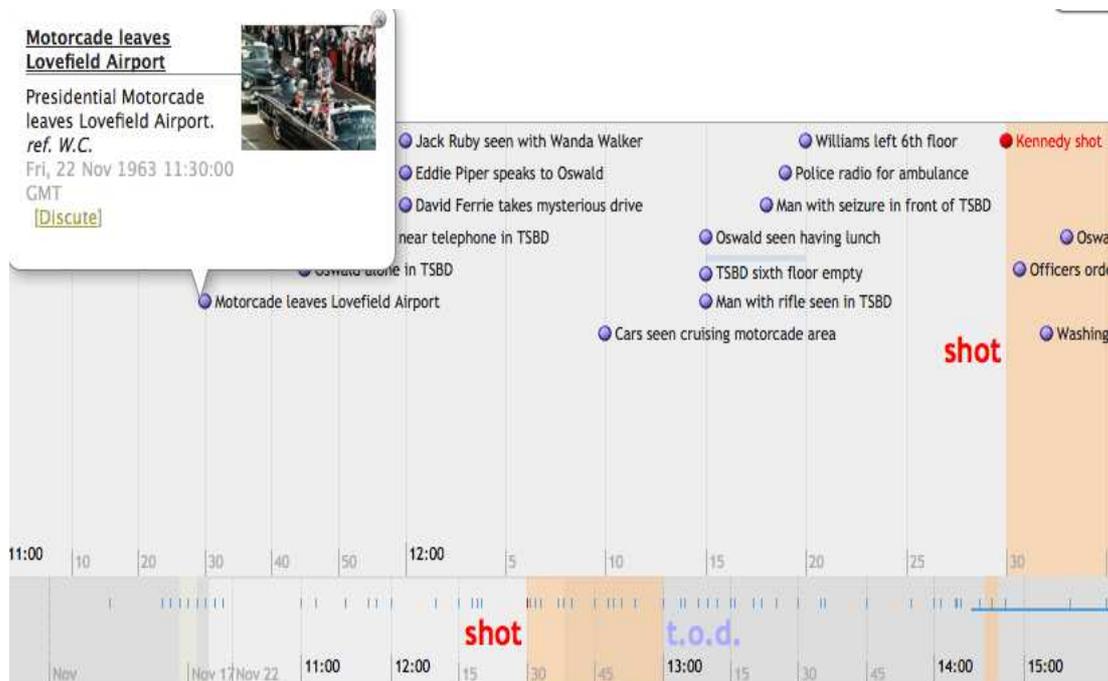


Figura 49. Navigazione semantica per *TIMELINE*

□ Con la **Navigazione *MAP INTERFACES*** (Figura 50) si può visualizzare, ad esempio:

- ✓ da dove è originata/o una certa pubblicazione, un progetto di ricerca;
- ✓ la rete di collaborazione di un ricercatore;
- ✓ temi di ricerca per aree geografiche;
- ✓ localizzazione di esperti di ricerca e le loro Università di origine;
- ✓ produzioni tematiche di ricerca delle istituzioni e così via.

L'utente può interagire con la mappa per restringere i risultati della ricerca [Clarke, 2008].

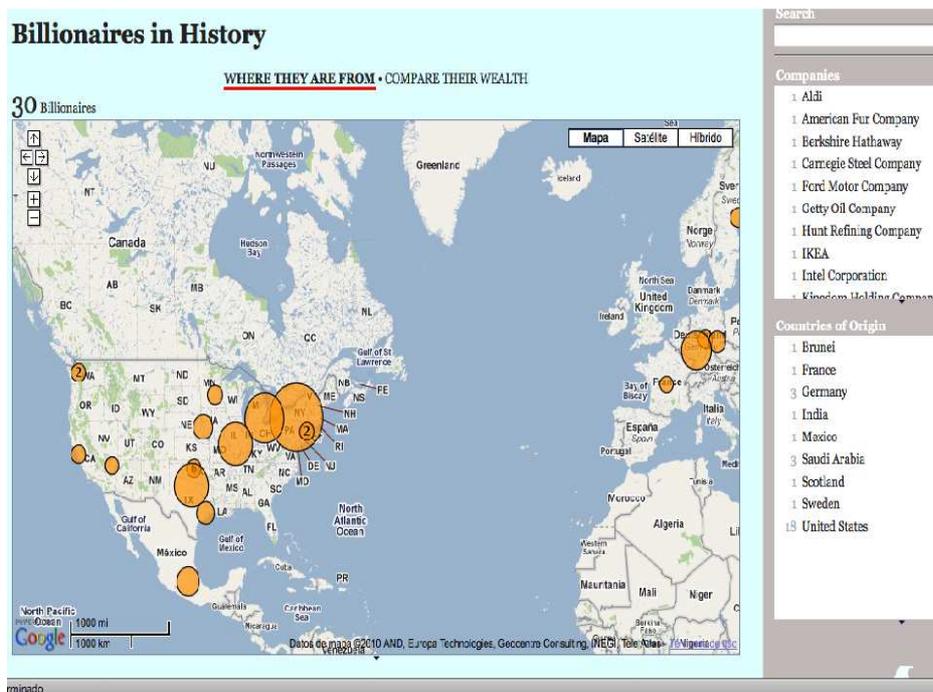


Figura 50. Navigazione semantica per *MAP INTERFACES*

Ricerca e Navigazione delle informazioni nelle pagine dei Profili di Autori depositanti

È possibile ricercare le risorse informative, visualizzare le loro statistiche d'uso, l'informazione bibliografica, i membri dei gruppi di ricerca di appartenenza direttamente dai Profili di Autori depositanti dell'Archivio.

[V. sito: "Archivio VOA3R: ricerca e navigazione nelle pagine dei Profili di Autori"] (Figura 51).



Figura 51. Archivio VOA3R: ricerca e navigazione nelle pagine dei Profili di Autori depositanti

Ricerca dei contenuti correlati in altri sistemi

Da un'interfaccia di ricerca di un Archivio [V. sito: Archivio arXiv.org] (Figura 52) è possibile consultare i contenuti correlati gestiti da altri Archivi, Sistemi editoriali.



Figura 52. Archivio arXiv.org: consultazione dei contenuti correlati gestiti dal Sistema editoriale *Public Knowledge Project*

“Web Design e NAVIGAZIONE dei contenuti: come facilitare la scoperta e l’accesso ai tuoi contenuti. Quello che ogni Esperto di Comunicazione deve sapere” (Tutorial): <http://www.masternewmedia.org/it/web-design-e-navigazione-dei-contenuti-come-facilitare-la-scoperta/>

Le strategie suggerite per la ricerca e navigazione dei contenuti digitali:

Ricerca dei Contenuti: PERTINENZA Vs. RICORRENZA Vs. DERIVAZIONE e SINONIMI.

Un forte indicatore della Pertinenza è se le parole che si cercano compaiono o meno nella serie di risultati, ma questo non significa che Pertinenza e Ricorrenza siano la stessa cosa. Un buon motore di ricerca va oltre la Ricorrenza, cercando Parole Derivate e tenendo conto dei Sinonimi.

Portare i lettori nel modo più veloce e più semplice possibile alla consultazione dei contenuti oggetto della loro ricerca.

Offrire la possibilità di consultare contenuti correlati o simili tramite i link/rimandi ad altri siti. Ad esempio, se una *Query* contiene nome di una persona, un link alla sua Biografia potrebbe essere il primo risultato mostrato.

Offrire la possibilità all’utente di abbonarsi agli *RSS (Really Simple Syndication) feed* tematici e alla *newsletter* per ricevere le notizie relative ad alcuni degli argomenti più approfonditi e consultati e ai nuovi contenuti nell’Archivio.

Ricerca dei Progetti nell’Archivio: <https://phaidra.cab.unipd.it/browse>

Caso d’uso: l’Archivio digitale PHAIDRA dell’Università di Padova.

4.8. Conclusioni: la sintesi del contesto del “Profilo Tecnico” di un Archivio Istituzionale

La valenza pratica degli aspetti costituenti il *Toolkit* di un Archivio Istituzionale (*Institutional Repository, IR*) deve costituire la vera chiave di lettura dei servizi di quest’ultimo, inteso come:

- 1) DEPOSITO di RISORSE digitali FACILMENTE FRUIBILI;
- 2) ARCHIVIO DIGITALE RAGIONATO e moderno basato su:
 - a) Buone Pratiche,
 - b) Protocolli di interoperabilità (quello di default è OAI-PMH) aventi l’obiettivo primario di garantire al sistema la massima interazione ovvero interoperabilità con altri sistemi digitali esterni, come aggregatori di Archivi Istituzionali e Disciplinari, piattaforme sviluppate entro i progetti nazionali e internazionali (es. Europea [V.sito]), motori di ricerca;
 - c) Modelli per la valorizzazione delle risorse a lungo termine tramite:
 - c.1) OAIS (*Open Archival Information System*), attualmente il modello concettuale di conservazione per eccellenza;
 - c.2) PREMIS (*PREservation Metadata: Implementation Strategies*);
 - c.3) gli identificatori persistenti come DOI, Handle, NBN;
 - c.4) procedure (es. migrazione, *refreshing*, emulazione) a supporto dei formati e l’ambiente macchina (*Software*) nel tempo;
- 3) un MOTORE di RICERCA potente e veloce che rende trasparente l’utilizzo del *Repository* e dei servizi ad esso connessi;
- 4) un INSIEME di interfacce *SOFTWARE* (prediligendo l’impiego di *Software Open Source* come *Fedora, Eprints, Dspace*) e SERVIZI (es. servizi modulari *SURplus-Open Archive, U-GOV-Ricerca*; Portali di Ateneo; piattaforme *e-Learning*) da usare:
 - a) come mattoncini di base per costruire funzionalità specifiche;
 - b) per erogare le risorse depositate;
 - c) per effettuare ricerche da altre applicazioni;
 - d) per mostrare una risorsa digitale collegata ad altre risorse; ad esempio una risorsa scientifica all’interno di un corso di *e-learning*, un antico codice digitalizzato all’interno un nuovo codice digitale e così via.
 - e) arricchire lo spettro dei servizi specializzati necessari alla realizzazione di specifiche e avanzate funzionalità a supporto di tutto il CICLO DI VITA di una risorsa digitale o digitalizzata.

Il Ciclo di vita di una risorsa comprende le seguenti fasi o macroprocessi di filiera (Tabella 13) che “accompagnano” la medesima durante il suo “viaggio” nel sistema.

1. IMMISSIONE <i>(INGESTION)</i>	Una procedura di caricamento nell’infrastruttura di risorse digitali (<i>born-digital</i>) o digitalizzati. Il processo di <i>INGESTION</i> è funzionale ai processi di Metadattazione e Catalogazione.
2. METADATTAZIONE	Un processo che segue la fase dell’analisi dei materiali esistenti e dei criteri decisi per la loro metadattazione, individuando dei Campioni rappresentativi delle diverse tipologie, a cui poter applicare degli specifici <i>Content Model</i> (modelli di contenuti informativi). I metadati <i>descrittivi, amministrativi e strutturali</i> (es. Dublin Core, MODS, METS, MAG) specifici per le tipologie di materiale devono: <ol style="list-style-type: none"> 1) essere codificati in conformità agli Standard più diffusi a livello internazionale;

2) arricchiti dai riferimenti a Vocabolari Controllati (terminologie, liste di interstazione per soggetto, liste di autorità dei nomi e dei luoghi geografici) secondo le esigenze locali del sistema e quelle della comunità di ricerca;

3) essere il più possibile interoperabili nell'ambito della gestione dei *Repository Istituzionali*:

3.1) l'interoperabilità tecnica tra gli Archivi avviene attraverso il protocollo l'OAI-PMH;

3.2) l'interoperabilità semantica dei contenuti si ottiene tramite:

3.2.1) l'impiego di Schemi di Classificazione, Thesauri;

3.2.2) tecnologie *Linked Data* (collegamento e il riuso dei dati web strutturati dello stesso contesto e ambito semantico);

3.2.3) tecniche per il *cross-linking* di *contenuti di ricerca* con i *set di dati scientifici* ad Accesso Aperto ed altri dati associati [V. sito": "*OpenAIREplus*"]

"I benefici che le scienze possono trarre dall'uso, il riuso e la combinazione di questi dati sono enormi in termini di *produttività, efficienza, risoluzione di problemi, di nuove scoperte*" [V. sito": "*Archeological Open Data*", p.15]

La preparazione dei Campioni con la *costruzione* dei *FILE* di METADATAZIONE deve essere realizzata facendo diverse PROVE di INGESTION fino al raggiungimento della forma più opportuna che permetterà al sistema di effettuare il CARICAMENTO del CAMPIONE e di attivarvi i servizi previsti per la sua disseminazione.

In parallelo deve essere sviluppato un *Software* specifico per la creazione di una PROCEDURA AUTOMATICA che crei, raccolga, integri e codifichi i metadati secondo il *CONTENT MODEL*, in conformità all'infrastruttura definita e in coerenza con i vincoli stabiliti dalla procedura di *Ingestion*.

I PACCHETTI di *INGESTION* - ovvero pacchetti di versamento (SIP: *SUBMISSION INFORMATION PACKAGE*) di soli Metadati e di Metadati e Oggetti digitali - vengono forniti nel sistema tramite file XML e in formato METS, strutturati secondo regole ben definite e forniti in modalità di:

1) *BULK* (i gruppi di file e i file di grandi dimensioni);

2) *DROP-BOX* (un Software multiplatforma *Cloud based* che offre un servizio di *file hosting* e sincronizzazione automatica di file tramite web);

3) *HARVESTING* (raccolta) di metadati tramite il protocollo OAI-PMH.

Queste procedure vengono elaborate da processi automatici in modo coordinato con tutte le altre componenti del servizio.

	<p>In seguito avviene il SALVATAGGIO del PACCHETTO di versamento nell'area di deposito (ARCHIVIAZIONE) del <i>Repository</i> tramite, ad esempio, i pacchetti AIP <i>ARCHIVE PACKAGE</i>.</p> <p>In particolare, dopo l'acquisizione di un pacchetto SIP - in base alla tipologia di entità specificata nello schema di metadati METS - vengono generati i pacchetti AIP che saranno oggetto di archiviazione nel <i>Repository</i>.</p>
<p>3. CATALOGAZIONE</p>	<p>Dopo l'impacchettamento automatico e deposito digitale tramite i pacchetti di versamento SIP - vengono creati i METADATI a livello COLLEZIONE, sia in termini <i>descrittivi</i> che <i>strutturali</i> per permettere l'attivazione dei servizi adeguati alle caratteristiche peculiari di ogni Collezione. I metadati a livello collezione sono le INFORMAZIONI di INDIRIZZAMENTO e di SINTESI DESCRITTIVA delle risorse, appartenenti alla Collezione specifica. Un buon esempio dello strumento <i>web-based</i> di CATALOGAZIONE è realizzato attraverso il CMS Open Source <i>DRUPAL</i> che permette di creare delle INTERFACCE di CATALOGAZIONE COMPLESSE, come:</p> <ol style="list-style-type: none"> 1) CATALOGAZIONE di METADATI descrittivi di Risorse e Collezioni digitali; 2) Definizione e utilizzo di VOCABOLARI CONTROLLATI. <p>Per la Soggettazione e il Controllo di autorità dei Soggetti delle Risorse possono essere adottati diversi Vocabolari Controllati. Ad esempio, a primo livello di Soggettazione può essere adottato un Thesaurus locale. Per un livello di soggettazione più approfondito potrebbe essere adottato il Thesaurus del NUOVO SOGGETTARIO di Firenze. I Vocabolari Controllati VID-SBN (identificativo di un Autore) e VIAF (<i>VIRTUAL INTERNATIONAL AUTHORITY FILE</i>) possono essere adottati anche per il controllo dei nomi di Autori. Il Thesaurus TGN (<i>THESAURUS OF GEOGRAPHICAL NAMES GETTY</i>) potrebbe essere adottato per la normalizzazione dei toponimi e così via;</p> <ol style="list-style-type: none"> 3) CLONAZIONE (duplicazione) di schede descrittive (di metadati); 4) caricamento (<i>UPLOAD</i>) e associazione di contenuti digitali alle risorse digitali; 5) VALIDAZIONE delle risorse digitali tramite un <i>workflow</i> "<i>Edit/Approve/Reject</i>"; 6) Associazione ad un <i>LOCAL IDENTIFIER</i>; 7) Esportazione verso un PACCHETTO di SIP. <p>Il servizio web della catalogazione prevista dal CMS Open Source <i>DRUPAL</i> permette inoltre la Catalogazione secondo gli Standard ISAD e ISAAR e prevede la gestione di seguenti tipologie di entità:</p> <ol style="list-style-type: none"> 1) RISORSE DIGITALI (un articolo, un'immagine, un libro, un video, un frammento audio, una cartografia, un oggetto 3D, un Software ecc.); 2) COLLEZIONI (una raccolta di risorse) e loro partizioni (serie,

	<p>sottoserie, raccolte, fascicoli, sottofascicoli ecc.);</p> <p>3) SOGGETTI VERSANTI (produttori) rappresentati da <i>Authority File</i> per la gestione normalizzata di loro denominazioni.</p> <p>Lo strumento web di Catalogazione realizzato attraverso il <i>CMS Open Source DRUPAL</i> pone particolare attenzione alla distinzione tra i <i>vari livelli di rappresentazione</i> di un oggetto digitale. Per evitare ogni ambiguità tra le risorse vengono inseriti i METADATI relativi alle seguenti entità:</p> <ol style="list-style-type: none"> 1) <i>Real Physical Object</i> (RPO) che rappresenta l'oggetto fisico (es. un libro, un edificio, un dipinto); 2) <i>Digital Representation Object</i> (DRO) che è un oggetto digitale ottenuto attraverso la digitalizzazione di un RPO; 3) <i>Digital Primary Object</i> (DPO) che è un oggetto "born digital".
<p>4. AMMINISTRAZIONE (<i>management</i>)</p>	<p>Il CONTROLLO AUTOMATICO dell'area di deposito digitale deve includere le funzioni di:</p> <ol style="list-style-type: none"> 1) Memorizzazione e gestione dei contenuti, nelle varie forme adatte alla conservazione e alla disseminazione. <p>Ogni OGGETTO caricato nel <i>Repository</i> implementato, ad esempio, sul <i>Software FEDORA</i> viene salvato sul <i>FILE SYSTEM</i> in una <i>DIRECTORY</i> che in base alla tipologia di datastream può essere gestita da <i>FEDORA COMMONS</i> stesso o indirizzata in modo diretto (<i>referenced datastream</i>);</p> <ol style="list-style-type: none"> 2) Memorizzazione e gestione dei metadati (descrittivi, tecnici amministrativi, strutturali). <p>Ad esempio, nel <i>Repository</i> implementato sul <i>Software Fedora</i> la sezione Metadati viene salvata in formato FOXML, mentre eventuali <i>file binari</i> presenti negli oggetti figli vengono rinominati e salvati in una speciale <i>gerarchia di directory</i> (o <i>datastore</i>) gestita da un algoritmo interno a <i>Fedora</i>;</p> <ol style="list-style-type: none"> 3) Memorizzazione e gestione dei derivati; 4) Memorizzazione e gestione delle informazioni di autorizzazione (diritti d'accesso); 5) Acquisizione automatica (<i>ingestion</i>) dei pacchetti SIP; 6) Controlli di integrità dei file; 7) Controllo Antivirus sui contenuti; 8) Integrazione con il processo di transcodifica (si pensi alla transcodifica generata dall'<i>ingestion</i>/versamento di un libro digitalizzato), tramite cui – in base al tipo di risorsa digitale in ingresso – vengono ottenute nuove forme (secondo le regole per ogni tipologia di elaborazione) di rappresentazione digitale utili alla disseminazione della risorsa come, ad esempio: <ol style="list-style-type: none"> 8.1) transcodifiche audio-video; 8.2) conversioni di immagini; 8.3) estrazioni di <i>thumbnail</i> ovvero un'anteprima di un'immagine più grande, che quindi viene presentata in formato ridotto; 8.4) transcodifiche OCR;

8.5) transcodifica dell'output generato da un altro servizio di trasformazione;

8.6) analizzatori della lingua;

8.7) estrattori di testo o analizzatori semantici.

Diverse transcodifiche possono essere ottenute tramite l'integrazione del *Software* del *Repository* con una serie di prodotti *Open Source* per la gestione, indicizzazione e trasformazione di contenuti digitali.

Per ottenere dall'oggetto digitale (*input*) il suo derivato (*output*) tramite il servizio di transcodifica – la piattaforma *Software FEDORA*, ad esempio, permette l'integrazione del sistema *Repository* con uno specifico set di API (*Application Programming Interface*) molto granulare implementato su un nodo dedicato per supportare l'intero *Workflow* di transcodifica.

L'API "CONCEPT MAPPER" (un *use case* dello strumento integrato nel *Digital Library Management System* del CINECA) - sfruttando la conoscenza disponibile in diverse risorse linguistiche (Wikipedia, EuroVoc, altre ontologie, dizionari e Thesaurus specialistici) - consente di *estrapolare* da un *documento testuale* i *concetti più rilevanti* in relazione al contesto e al dominio del documento stesso. Al termine dell'elaborazione viene generato un *File RDF SKOS* (all'interno di metadati semantici) che può essere inserito in un *datastream* del *Software FEDORA* e indicizzato per la ricerca dei contenuti;

9) *Indicizzazione* delle risorse e delle collezioni per alimentare il processo di ricerca (*Search*).

L'indicizzazione automatica può avvenire, ad esempio, tramite il motore di ricerca integrato *Open Source Apache SOLR*, riguardando:

9.1) i *metadati* descrittivi;

9.2) le *relazioni* tra gli oggetti presenti nel *Repository*;

9.3) i *contenuti testuali originali* e *derivati* da *trascrizioni automatiche* o da *analisi semantiche*,

mantenendo aggiornato l'Indice dei contenuti interrogabile dalle applicazioni *front-end*.

In caso di eliminazione di un oggetto, il sistema genera un opportuno messaggio in modo che il motore di ricerca proceda con l'eliminazione dall'indice;

10) Creazione automatica di *oggetti digitali derivati*, adatti alle forme di disseminazione previste dal servizio (es. immagini a risoluzioni minori, *Book viewer*, codifica di filmati per lo *streaming*);

11) Arricchimento (*enrichment*) delle risorse digitali con:

11.1) *generazione automatica* di *metadati semantici* ovvero l'affiancamento delle proprietà di metadati descrittivi con valori semantici di Classificazioni, Thesauri;

11.2) l'aggiunta di informazioni ottenute tramite estrazione di testo (OCR);

11.3) l'identificazione della lingua di un testo;

12) Archiviazione di metadati e contenuti delle risorse;

	<p>13) Controllo delle modifiche effettuate sugli oggetti, compreso il monitoraggio dei processi di transcodifica, che devono essere registrate in uno specifico <i>datasteam</i> di sistema non modificabile (caso d'uso: il <i>Software Fedora</i>);</p> <p>14) Gestione dei Diritti di proprietà intellettuale (di <i>Copyright</i>: rapporti Autore-Editore; Autore-Istituzione; Autori-Lettori) e di Licenza associata [Shearer et al., 2012] alle risorse digitali costituenti l'oggetto informativo.</p> <p>Per supportare tale processo in un Archivio Istituzionale è opportuno integrare il servizio SHERPA/RoMEO [V. sito] che ha sviluppato uno strumento "<i>Rights on METadata for Open archiving</i>"- Politiche dell'impiego (deposito, disseminazione ecc.) delle pubblicazioni, adottate da diversi editori.</p> <p>15) Gestione degli accessi e delle autorizzazioni basata su considerazioni e sull'analisi dello stato dell'arte dei sistemi di gestione delle identità (definizione delle Classi e dei Ruoli di utenza) e di autenticazione attualmente utilizzati in rete.</p> <p>Ad esempio, nel <i>Software Open Source Fedora Commons</i> le autorizzazioni vengono inserite e memorizzate in qualsiasi oggetto ed espresse secondo lo standard XACML. XACML consente un livello di granularità molto dettagliato nella definizione delle <i>Policy</i> sia a livello di utente, sia di azione sia di risorsa;</p> <p>16) <i>backup</i> quotidiano dei contenuti.</p>
<p>5. DISSEMINAZIONE di collezioni e risorse</p>	<p>La disseminazione è basata su:</p> <ol style="list-style-type: none"> 1) navigazione; 2) ricerca; 3) visualizzazione delle risorse; 4) autenticazione e autorizzazione degli accessi alle risorse digitali. <p>1) Navigazione (<i>browsing</i>) nell'Archivio può avvenire per:</p> <ol style="list-style-type: none"> 1.1) GERARCHIE di strutture (organizzative, collezioni e tipologie di contenuti); 1.2) informazioni contenute nei METADATI descrittivi; 1.3) filtri a FACCETTE. <p>La CLASSIFICAZIONE "a FACCETTE" permette di catalogare un singolo <i>asset</i> digitale sotto più categorie (le faccette), ciascuna descrittiva (anche tramite i metadati) di un aspetto (faccia) particolare dell'assetto.</p> <p>La selezione di una FACCETTA aggiunge un FILTRO A CASCATA che restringe i risultati di ricerca in quella categoria ovvero un clic su una faccetta serve come filtro istantaneo in una schermata di <i>browsing</i>.</p> <p>Le entità implicate nelle relazioni gerarchiche possono essere organizzate tramite l'uso di relazioni espresse secondo il modello RDF (<i>Resource Description Framework</i>) "<i>Subject-Object-Value</i>". Queste stesse relazioni possono essere utilizzate per modellare aggregazioni <i>ad-hoc</i> e personalizzate di entità.</p> <p>2) Ricerca (<i>Advanced</i> e <i>Simple Search</i>) può avvenire tramite:</p>

- 2.1) metadati indicizzati;
- 2.2) set di keyword minimale;
- 2.3) tipologie di risorse di interesse;
- 2.4) campo da utilizzare per l'ordinamento;
- 2.5) *full-text* all'interno dei testi.

Alcuni motori di ricerca integrati – come ad esempio *Open Source Apache SOLR* – possono individuare le risorse “vicine“ (correlate) rispetto ai metadati descrittivi delle risorse e delle collezioni.

La ricerca *full-text* può essere applicata ai contenuti testuali o alla trascrizione di in testo di immagini derivanti da processi di digitalizzazione quando disponibile.

Le informazioni contenute nell'indice del motore di ricerca sono costantemente aggiornate rispetto alle informazioni che sono presenti nel *Repository* (es. aggiunta, modifica o eliminazione di un oggetto);

3) Visualizzazione può avvenire per tipo di risorsa sotto forma della sua rappresentazione digitale più adatta:

- 3.1) come API a granularità fine con cui i programmatori possono realizzare le applicazioni di visualizzazione specifiche;
- 3.2) come *building blocks* integrati con le piattaforme come, ad esempio, U-GOV Ricerca, SURplus *Open Archive*, Moodle (*Learning Management System Open Source*), *DRUPAL*, con cui si potrebbe costruire uno specifico Portale o applicazione web.

L'integrazione di tutti i servizi coinvolti nel funzionamento del sistema Archivio Istituzionale e la loro esposizione verso il livello applicativo del medesimo potrebbe essere regolato attraverso una piattaforma API scelta per fungere da *unico punto di accesso* verso tutti i servizi dell'Archivio. La piattaforma API o una serie di API devono essere invocabili da una qualsiasi applicazione client (es. il Portale).

Le API possono essere implementate nell'Archivio attraverso il servizio REST (un *web service* che propone una visione del Web incentrata sul *concetto di risorsa*) e possono essere ricondotte a due tipologie distinte:

- a) API per il recupero del contenuto degli oggetti che forniscono generalmente una risposta in forma binaria (es. un'immagine, un pdf, un video ecc.);
- b) API per la navigazione dei metadati;

Lo sviluppo di ricchi strumenti di disseminazione, semanticamente e interattivamente appropriati, per la ricerca, la navigazione (*browsing*), il recupero e la visualizzazione di un determinato tipo di oggetto digitale, deve essere definito all'interno del Modello (*Content Model*) della tipologia cui appartiene l'oggetto (es. produzione scientifica digitale, libri antichi e moderni digitalizzati, immagini in formato digitale, materiale audiovisivo, materiale audio, materiali didattici).

Al fine di un'efficace disseminazione di contenuti di un Archivio e' opportuno, inoltre, sviluppare un'efficiente strutturazione dei servizi di:

	<p>a) <i>reference</i> (attività di supporto agli autori, FAQ); b) <i>feedback</i> per l'interazione attiva con l'utente; c) collegamento dell'Archivio con i servizi esterni/<i>outsourcing</i> (entro e fuori dell'Ateneo) per promuovere la conoscenza sull'Archivio aumentando la sua visibilità in rete a livello globale.</p>
<p>5. Programmazione di Long Term Digital PRESERVATION e di ACCESSO granulare (secondo le condizioni d'accesso)</p>	<p>La Programmazione di <i>Long Term Digital PRESERVATION</i> e di ACCESSO ai contenuti digitali deve essere in linea con le migliori pratiche derivanti dalla conservazione a lungo termine:</p> <ol style="list-style-type: none"> 1) conformità al modello dal modello internazionale a pacchetti OAIS (<i>Open Archival Information System</i>, ISO 14721:2012); 2) implementazione dello schema di metadati di conservazione PREMIS; 3) l'integrazione con i più diffusi sistemi di <i>Persistent Identifier</i>. <p>Per quanto riguarda l'archiviazione e la conservazione dei contenuti digitali, la scelta del <i>Software Open Source FEDORA Commons</i> è indubbiamente strategia e vincolante ai fini della gestione a lungo termine dei materiali. <i>FEDORA Commons</i> [V. l'indagine dell'Archivio PHAIDRA, PARTE II, Cap.3] fornisce un ambiente sicuro, articolato ed orientato non solo alle integrazioni con Web Services e tecnologie del Web Semantico (l'utilizzo delle relazioni RDF), ma anche alla Conservazione a lungo termine. <i>Fedora</i> permette anche l'applicazione di diversi <i>Content Model</i> che sono più rispondenti alle esigenze di gestione e conservazione delle peculiarità di diverse tipologie di oggetti digitali.</p>

Tabella 13. Fasi o macroprocessi di filiera coinvolti nel Ciclo di vita di una risorsa in un Archivio digitale

I macroprocessi appena descritti in Tabella 13 presentano il sunto degli aspetti costitutivi del *Toolkit* passato in rassegna nel presente Capitolo, da:

- ✓ conferire come il *Workflow* delle soluzioni prototipali al sistema opportunamente costruito;
- ✓ integrare nell'ipotesi di progettazione e di future estensioni del servizio Archivio Istituzionale esistente a nuovi contesti ai quali un eventuale Progetto di quest'ultimo potrebbe far fronte.

L'attenzione del prossimo Capitolo sarà focalizzata sul processo di METADATAZIONE, particolarmente sullo sviluppo di un profilo di metadati che potrebbe essere applicato alla descrizione delle risorse digitali in un Archivio Istituzionale. I metadati scelti per tale compito devono corrispondere ai criteri di qualità, quindi essere sviluppati in conformità agli Standard più diffusi a livello nazionale e internazionale e arricchiti dai riferimenti a Vocabolari Controllati (es. specifici Thesauri) e a Classificazioni – preferibilmente con i valori non-letterali, basati sugli identificatori web < http://URI> - laddove possibile e necessario. E' di vitale importanza che le risorse dell'Archivio interagiscono tra loro e con le risorse di altri sistemi - in base al contesto e affinità semantica del loro contenuto - per cui devono essere descritte anche tramite i metadati che definiscono le loro possibili relazioni (es. "*isMemberOF*", "*isPartOF*", "*isMemberOfCollection*").

Il profilo di metadati che sarà presentato nel prossimo Capitolo viene ispirato, per uno sviluppo di una adeguata qualità di metadattazione per rappresentare e disseminare con l'efficienza i contenuti, le proprietà e le relazioni delle risorse digitali di un Archivio da:

- ✓ “Linee guida per la creazione e la gestione di metadati nei *Repository* Istituzionali“ [V.sito] approvate dalla Commissione Biblioteche della CRUI (Italia) nel 2012;
- ✓ le Raccomandazioni “LODE-BD“ (*Linked Open Data-enabled Bibliographical Data*) [V.sito] che riportano il metodo su come appropriatamente selezionare le strategie di codifica per la produzione dei dati bibliografici abilitati come *Linked Open Data (LOD)*. Le Raccomandazioni “LODE-BD“ si stanno promovendo dal 2012 tramite le iniziative della Comunità online AIMS (*Agricultural Information Management Standards*);
- ✓ altri principi e strategie.

PARTE III

Potenziamento dei metadati dalle strategie "LODE-BD" (*Linked Open Data – enabled Bibliographic Data*) per una migliore comunicazione dei contenuti di ricerca

CAPITOLO 5

Sviluppo di uno schema di metadati a supporto descrittivo dei contenuti di ricerca negli Archivi Istituzionali alla luce delle strategie "LODE-BD" (*Linked Open Data – enabled Bibliographic Data*)

5.0. Introduzione

Questioni come:

1) modellazione e adeguamento dei metadati (dati strutturati che descrivono altri dati) agli Standard internazionali ampiamente riconosciuti e utilizzati, per favorire una maggiore interoperabilità tra le applicazioni;

2) organizzazione dei metadati tramite il loro incrocio in modo coerente con diversi Schemi e Profili applicativi dei metadati, con gli strumenti di controllo di autorità (*Authority Control*) e schemi semantici come Classificazioni, Ontologie, contribuendo all'armonizzazione, alla normalizzazione e semantizzazione dei valori di metadati;

3) attività di *Licensing* (scelta e applicazione di Licenze) e di Conservazione a lungo termine (applicazione dello Standard OAIS: *Open Archival Information System*) valorizzando lo scopo e l'utilizzo di risorse digitali;

4) pubblicazione e abilitazione dei contenuti digitali (risorse e metadati) in modalità *Linked (Open) Data* - in quanto questa modalità si sta dichiarando utile, fattibile e applicabile a tutte le tipologie e formati di dati [V. siti: *Global Interoperability and Linked Data in Libraries; Open Repositories 2013*]

stanno diventando per i gestori di diversi Archivi Istituzionali sempre più attuali e urgenti.

Le questioni appena elencate creano un terreno fertile per il proliferare dei modelli e delle attività che prendono sempre più piede nel mondo degli Archivi di ricerca digitali, tra cui:

- Archivi digitali dei Dati di Ricerca (*Research Data Archive*) e *Data Management Plans* per programmare l'intero ciclo di vita dei contenuti digitali [V. sito: *Open Repositories 2013; Data Management Plans*];
- interoperabilità tra le piattaforme compatibili con le infrastrutture OAI (*Open Archive Initiative*), per catturare, aggregare e ricercare in modo trasversale sul Web i contenuti di ricerca risultanti come l'*output* di Data e Service Provider OAI [V. siti: *Open Archive Initiative; European Interoperability Framework*];
- interoperabilità tra gli Archivi digitali di ricerca e CRIS (*Current Research Information Systems*) per contribuire allo sviluppo dell'Anagrafe locale e nazionale delle ricerche, di servizi e applicazioni esterne [Cassella, Bozzarelli, 2011; V. sito: *Rome Declaration on CRIS and OAR*];
- gestione di metadati in modo transdisciplinare, con ampio riutilizzo dei dati e servizi di Archivi di ricerca;
- necessità della creazione dei servizi a valore aggiunto affidabili e duraturi sopra gli Archivi aventi i requisiti di *Trusted Digital Repositories* [V.sito].

In questo ricco contesto di attività e iniziative una dovuta attenzione deve essere prestata allo sviluppo di Metadati di qualità a supporto descrittivo, strutturale e gestionale dei contenuti negli Archivi digitali di ricerca, in conformità con le norme e Standard vigenti, Linee Guida, Schemi e

Profili di applicazione di metadati ampiamente in uso e *Best practice* riferiti alle strategie di codifica di valori dei metadati.

Il presente Capitolo è incentrato sulla questione dell'*abilitazione* dei *metadati descrittivi* dei contenuti di ricerca degli Archivi Istituzionali per lo spazio *dei dati interconnessi nel Web Semantico*, ovvero l'*abilitazione* dei metadati proposti dalle "Linee Guida per la creazione e la gestione di metadati nei Repository Istituzionali" [V.sito] come *Linked Data* tramite le strategie di codifica proposte dalle Raccomandazioni "*Linked Open Data-enabled Bibliographical Data*" [V.sito].

In questo modo i metadati dell'Archivio possono acquisire i requisiti necessari per entrare nel reticolo dei dati interconnessi del *Web di Dati* (Web 3.0) basandosi sul valore delle relazioni proiettate sul contesto e sulla semantica dei diversi *dataset* (pacchetti di contenuti) presentati graficamente nello spazio *Linking Open Data (LOD) Cloud* [V.sito; APPENDICE A].

La nuvola *LOD Cloud* fornisce un panorama di una diversità dei pacchetti di dati (*dataset*) pubblicati online per mezzo della tecnologia *Linked (Open) Data* e copre attualmente più di 50 miliardi dei "fatti" (contenuti, dati) relativi a diversi domini di conoscenza [V.sito: *Best practice Report on cultural heritage*]. I contenuti pubblicati come *Linking Data* e "incapsulati" nei pacchetti di dati *Open* - ovvero pubblicati sotto le Licenze Aperte - possono essere:

- facilmente interconnessi con altri pacchetti di dati pubblicati sul Web come *Linked (Open) Data*;
- ri-utilizzati (consumati) e
- arricchiti nel loro aspetto semantico tramite le connessioni opportunamente create da diversi soggetti produttori che operano con i loro contenuti sul Web.

In questo modo si contribuisce alla creazione di una comunicazione innovativa sul web basata sui sempre più nuovi reticoli della conoscenza.

Negli ultimi anni, anche diversi *dataset* bibliografici - tra cui risorse digitali e loro collezioni, metadati, *Authority File* (es. vocabolari controllati mono- multi-lingue, Terminologie) e Schemi semantici (es. Classificazioni, Thesauri) sono stati pubblicati e riutilizzati sul Web in base allo "Schema di distribuzione dati sul Web a 5 stelle", secondo i principi di Tim Berners-Lee, il creatore di World Wide Web [Tim Berners-Lee, 2006; V. siti: "Tim's 5 star" *Open Data plan with examples; Linked Data for Libraries*].

Il Rapporto *Library Linked Data (LLD)* e la sezione del Registro CKAN (*Comprehensive Knowledge Archive Network*) per *Library Linked Data* [V. siti], il servizio online *Linked Open Vocabularies (LOV)* [V. sito] che contiene l'elenco delle Ontologie Web RDFS, il Seminario internazionale "*Global interoperability and Linked Data in libraries*" [V. sito] e altre iniziative [APPENDICE A, §1.2.1.1] possono essere citati entro i primi più importanti "testimoni" che riportano e descrivono la proliferazione delle attività connesse con la pubblicazione di *Linked Data* su scala globale in rete.

Un panorama delle applicazioni che consumano *Linked Data* come, ad esempio, *faceted browsers*, *Linked Data browsing*, motore di ricerca *Linked Data*, "*On-the-fly*" *mashups* sono stati ben descritti nel documento "*Consuming Linked Data*" [Sequeda, 2011; V. sito: *Faceted Application of Subject Terminology*].

Per tradurre il successo iniziale di *Linked (Open) Data* in una realtà stabile all'interno dell'universo bibliografico, considerando anche le loro possibili connessioni con il *Social Web* o Web 2.0 e con i dati commerciali, ci sono ancora diverse questioni da affrontare:

- la necessità di alfabetizzazione su *Linked (Open) Data*, particolarmente la promozione della conoscenza sui benefici di pubblicazione, riutilizzo e integrazione delle risorse bibliografiche come *Linked (Open) Data*, direttamente - attraverso gli Standard e indirettamente - attraverso la promozione delle "buone pratiche";
- devono essere ancora rivalutati i vari requisiti per esprimere modelli di progettazione di

metadati - sia come modelli dei contenuti per i formati “*Linked-Data-compatibili*” che come gli elementi chiave per la creazione e consumo di metadati coerenti entro diverse comunità di pratica, secondo una struttura comune per lo scambio di dati Web a livello internazionale - *Resource Description Framework* (RDF) [V. sito: *Linking to the Future*];

- vi è la necessità di un’ampia promozione e condivisione pratica delle strategie di codifica di metadati per abilitare (*enable*) i loro valori come *Linked Data*. Un buon prospetto di tale codifica è fornito dalle Raccomandazioni “LODE-BD” [APPENDICE A, § 2.6.9] [De Robbio e Giacomazzi, 2011];
- devono essere ancora meglio esplorati, formalizzati e condivisi - come modelli comuni tra diverse comunità di pratica - i processi di *allineamento automatico* tra gli elementi di metadati e dataset *Linked Data - enabling*;
- è importante continuare ad indagare e promuovere il concetto di *trust* (affidabilità) riferito ai *Linked Data*. Solo i modelli di dati affidabili devono essere pubblicati come *Linked Data*. Il Progetto LOD2 [V. sito] del *FP7 Information and Communication Technologies Work Programme* sviluppa gli strumenti per la ricerca, navigazione e controllo di affidabilità di *Linked Data*;
- è opportuno condividere tra varie comunità di ricerca i modelli di pubblicazione dei dati scientifici sulla cima di *Linked Data* [Bechhofer, 2013, p.5] già presenti nello spazio di *Linked (Open) Data Cloud*;
- deve essere valorizzato e promosso il ruolo degli Archivi Istituzionali, sia in veste degli editori delle proprie risorse¹ come *Linked (Open) Data* che come implementatori di *Linked (Open) Data* che possono arricchire il contesto e la semantica delle medesime risorse (es. l’impiego di Thesauri, Classificazioni, risorse di altri Archivi e piattaforme pubblicate come *Linked Open Data*). Per esempio, nel *Repository Istituzionale* dell’Università di Oregon - “ScholarsArchive @ OSU” [V. PARTE II, Cap. 3] le Tesi e le Dissertazioni depositate sono state pubblicate come dataset *Linked Data*, dopo aver convertito i metadati descrittivi MARC e Dublin Core nel modello dei dati strutturati RDF [APPENDICE A, § 1.2.1] e attribuito per ogni “Risorsa”, “Autore” e “Soggetto” gli identificatori persistenti web URIs basati sullo standard *Handle*. Inoltre sono stati sviluppati i collegamenti (*link*) dai dataset espressi come *Linked Data* ai dataset *Linked Data* esterni [Johnson, Boock, 2012, pp. 6, 12]. L’interrogazione dell’intero dataset *Linked Data* dell’Archivio è possibile tramite il protocollo SPARQL (*Protocol e RDF Query Language*) che può interrogare i dati nel database della piattaforma Software per le triple RDF -*TripleStore*.

5.1. Perché *Linked Data*

Linked Data [V. APPENDICE A, §1.2.1.1] si riferisce ad un insieme di *Best Practice* per la pubblicazione e la connessione dei dati strutturati (tramite il modello di dati RDF) sul Web. Dal 2006 [Tim Berners-Lee, 2006] il concetto ‘*Linked Data*’ si è evoluto come candidato molto promettente e altamente raccomandato dalla comunità Web per affrontare una delle più grandi sfide nel settore della gestione delle informazioni intelligenti - l’uso del *Web* come piattaforma per l’interconnessione e l’integrazione di dati e di informazioni, oltre alla interconnessione e ricerca dei documenti ipertestuali.

¹ Nel contesto IR il termine “risorsa” può denotare un articolo, monografia, tesi, intervento alla conferenza, rapporto di ricerca, materiale di presentazione, learning object, dati di ricerca, metadati che li descrivono.

Il concetto *Linked Data* (LD) è stato formalizzato all'interno del già citato schema "*Tim's 5 star*" [V. sito], i cui principi sono stati riassunti tramite il riferimento a cinque stelle, come segue (Figura 53):

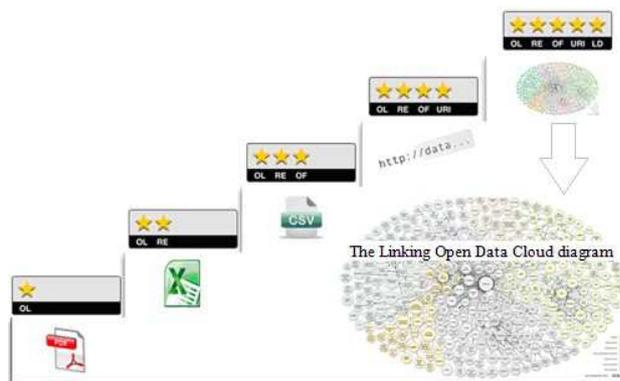


Figura 53. Schema "*Tim's 5 star*" di distribuzione di *Linked Open Data* a 5 stelle

- ☆ - si riferisce alla necessità di rendere sul Web i *dataset* (contenuti, pacchetti di dati) disponibili in qualsiasi formato tramite una Licenza Aperta;
- ☆☆ - si riferiscono alla necessità di rendere sul Web i *dataset* come dati strutturati attraverso il modello dei dati RDF;
- ☆☆☆ - si riferiscono alla necessità dell'utilizzo dei formati non proprietari (ad esempio CSV invece di Excel);
- ☆☆☆☆ - si riferiscono alla necessità di identificazione dei *dataset* sul Web tramite gli indirizzi web URIs, in modo che diversi agenti Web possono puntare a questi indirizzi e creare una varietà di collegamenti (*link*) basati su URIs;
- ☆☆☆☆☆ - si riferiscono alla necessità di collegare i *dataset* di partenza con altri *dataset* tramite gli identificatori Web dello schema globale URI, per creare il contesto entro lo spazio *Linked Open Data Cloud*.

Linked Data non significa, quindi, solo pubblicare i *dataset* sul Web. In primo luogo, *Linked Data* significa creare connessioni, così che diversi agenti (persone e sistemi web) possano esplorare una varietà di reti di conoscenza creati sui dati interconnessi.

Con *Linked Open Data*, quando si dispone quindi di alcuni dati di partenza *Open* - ovvero pubblicati tramite Licenze e Formati aperti - si possono creare diversi dati collegati, entrando apertamente nell'esplorazione dello spazio globale del Web dei Dati, *Linked Data*, o Web 3.0 – considerato un'espressione pratica del Web Semantico.

Linked Data possono essere quindi sia aperti (*Open*) che chiusi (dati che possono essere usati per esigenze strettamente locali, o dati cui Licenza dell'utilizzo deve essere comprata).

Linked Open Data (LOD) si riferiscono solo ai dati aperti (serie di dati pubblicati con Licenze *Open*) [V. sito: *Open Data Handbook*; Berners-Lee, 2006].

Il paradigma *Linked Open Data* è una strategia *Linked Data* per l'identità globale [Glaser, Halpin, 2012, p.9] di *Open Data* tale da permettere riutilizzo, liberando, condividendo e arricchendo i dati sul Web.

Qui di seguito vengono elencati alcuni benefici che possono derivare dalla pubblicazione e/o l'allineamento delle risorse digitali degli Archivi Istituzionali secondo le strategie *Linked Data* [V. siti: *Benefits of the Linked Data Approach*; *LOD-Around-The-Clock*].

Linked Data (LD) consentono di:

1. collegare, condividere, e interrogare (meta)dati provenienti da diverse fonti e formati. LD ha l'obiettivo di organizzare un ambiente di "silos" di dataset disconnesse in uno spazio di dati (contenuti) strutturati e collegati;
2. evitare la ridondanza (duplicazione) dei dati, mantenendoli inoltre aggiornati;
3. creare riferimenti incrociati ai dati di autorità e file semantici disponibili come LOD;
4. creare il *bookmarking* (segnalibri) nei domini informativi LOD incrociati della conoscenza enciclopedica come, ad esempio *DBpedia*, *Open Library*, *Freebase*, *GeoNames*, per riutilizzare i loro pezzi di informazione al fine di arricchire il contesto semantico delle risorse dell'Archivio;
5. elaborare direttamente i dati senza essere confinati dalle prestazioni di alcun *Software* particolare;
6. effettuare l'aggregazione dei dati, i calcoli, la visualizzazione, l'accesso, l'esportazione, controllo granulare sopra gli elementi di dati (ad esempio il bilanciamento del caricamento dei dati, *caching*²);
7. migliorare la valutazione dei contributi nelle discipline di ricerca specifiche. Strategie di LD possono riunire i dataset delle stesse discipline ospitati nei *Repository* disparati di tutto il mondo,
8. attrarre il nuovo pubblico di utenti dai ricchi contenuti sviluppati su *dataset* LOD degli Archivi Istituzionali, tramite l'impiego di API (*Application Programming Interfaces*) e *Web mash-ups* che combinano spesso API "generali" [Jarrar, Dikaiakos, 2012];
9. migliorare l'esperienza di ricerca e navigazione dell'utente, basata su insiemi di dati collegati e contestualmente pertinenti. Gli utenti saranno più propensi a visitare il sito di un certo Archivio o Portale potenziato dalle strategie di pubblicazione dei contenuti offerti dalla tecnologia *Linked Data*.

5.2. Importanza dell'uso di vocabolari controllati e schemi semantici

Vocabolari controllati e schemi semantici – come, ad esempio, le Liste di controllo di autorità dei Nomi in forma standard, Sistemi di classificazione, Thesauri, Mappe Topiche, Ontologie - sono noti genericamente come Sistemi di Organizzazione della Conoscenza (*Knowledge Organisation Systems, KOSs*) [V. PARTE II, Cap. 4, §§ 4.3, 4.4].

Questi Sistemi, tramite le voci di autorità, sinonimi, varianti ortografiche, gerarchie tesaurale, strumenti concettuali analoghi (classi, associazioni), offrono un modo sistematico per organizzare, normalizzare, arricchire semanticamente, potenziare l'accesso e il recupero dei contenuti digitali [Guerrini, Tillett, Sardo, 2003, p.11]. Senza affinare la descrizione delle risorse digitali dai valori di KOSs, sia gli utenti che i sistemi che ospitano le medesime risorse saranno ostacolati nei loro sforzi per accederle e aggregarle (secondo le proprietà comuni) [Salo, 2009].

Diversi KOS sono mantenuti da una Autorità (es. NACO Autorità della *Library of Congress* [V. sito]) che assicura che tutti i termini siano definiti in modo coerente e abbiano relazioni ben definite. In teoria, qualsiasi pezzo di informazione è suscettibile al controllo d'autorità - come ad esempio Nomi personali e di Enti, Titoli uniformi, Serie e Soggetti.

Intestazioni uniche (termini espressi da una stringa o un identificatore dell'indirizzo web) - assegnate, per esempio, a pezzi di informazione come "Autore", "Soggetto", "Titolo" e utilizzate poi in modo coerente, univoco, e senza ambiguità per descrivere tutti i riferimenti in relazione ai medesimi – possono popolare un database chiamato File di autorità (*Authority File*).

Questi File, così come i loro "legami logici" ad altri file collegati, devono essere mantenuti e aggiornati dai professionisti di metadati e altri professionisti dell'informazione.

² La memoria *cache* è una memoria informatica temporanea, non visibile al programmatore, che memorizza un insieme di dati che possano successivamente essere velocemente recuperati su richiesta.

Diversi vocabolari controllati e semantici sono stati espressi con i formalismi SKOS (*Simple Knowledge Organization Systems*) [V. sito] e pubblicati come LOD dataset [V. sito: *Library Linked Data Incubator Group: Datasets, Value Vocabularies*], in modo da essere comprensibili, condivisi e riutilizzati sul Web. L'uso dei KOS (disponibili online come LOD) per qualificare determinati valori dei metadati negli Archivi Istituzionali potrebbe, a sua volta, abilitarli (*enable*) come *Linked Data*, così come, ad esempio, nei Repository *Open Access* “VOA3R” e “ScholarsArchive@OSU” [V. PARTE II, Cap. 3].

Nonostante una grande offerta di KOS disponibili come *Linked Open Data* (quindi disponibili per un loro ri-uso aperto sul Web, senza alcuni restrizioni), per una loro consistente implementazione e utilità negli Archivi Istituzionali vi sono ancora alcune questioni aperte da affrontare. Ad esempio, non ci sono KOS che possano essere comunemente applicabili in tutti gli Archivi al fine di qualificare/normalizzare i Nomi propri degli Autori delle Tesi così come di tanti altri Autori che depositano i loro lavori negli Archivi Istituzionali. A proposito, l'impiego del file VIAF (*Virtual International Authority File*) disponibile sul web come *Linked Open Data* per qualificazione dei diversi Nomi degli Autori potrebbe essere molto limitato a questo scopo. In questo caso serve:

- sviluppare i database locali con le voci di autorità identificate con gli indirizzi web persistenti URI, al fine di contribuire ad una risoluzione univoca del controllo dei Nomi sul web;
- pubblicare le entità di tali database come *Linked Data* e distribuirli sotto *Licenze Aperte*, che consentirebbe ad altri agenti web di conoscerli e di creare i derivati (es. versioni multilingue, collegamenti con altri database di autorità e file semantici disponibili in *Linked Open Data*).

L'applicazione alla descrizione delle risorse informative degli Archivi dei valori di Vocabolari controllati e Schemi di classificazione standard è più costosa che l'assegnazione di semplici parole chiave non controllate. Ciononostante, investimenti in tal senso spesso risultano di molto vantaggio [V. sito: *NISO Framework Working Group*] sia per l'organizzazione delle risorse che per gli utenti d'Archivio in vista di una migliore ricerca, navigazione (es. navigazione e ricerca dinamica tramite la voce “Soggetto” per mezzo delle faccette basate sulle voci dei Thesauri, delle Classificazioni disponibili sul Web come *Linked Data*) e accesso [V. PARTE II, Cap.4, § 4.7].

In Tabella 14 vengono presentati alcuni Vocabolari Controllati e Schemi Semantici (Classificazioni, Ontologie) cui entità costitutive sono identificate con gli indirizzi web – *<http://URIs>*, quindi con i valori “non letterali” riferiti a quelli letterali (stringe di Soggetto, Classi, voci di Thesauri). L'implementazione di tali KOS nell'Archivio rappresenta non solo un approccio innovativo per il controllo di autorità e per l'arricchimento semantico delle proprietà dei metadati, ma favorisce anche la creazione degli “accessi multifocali” ovvero degli accessi unici proiettati a una diversità di risorse distribuite in modo trasversale sul web ma aventi le unità informative del contenuto in comune.

Vocabolario o Schema per la codifica delle proprietà di metadati	Descrizione
“RESPONSIBLE BODY” e valori di metadati correlati	
Virtual International Authority (VIAF): < http://viaf.org/ >	È un Progetto congiunto di diverse Biblioteche nazionali più alcune Agenzie di biblioteche regionale e trans-nazionali, sostenuto dall'OCLC (<i>Online Computer Library Center</i>). L'obiettivo del file VIAF è quello di ridurre i costi e aumentare l'utilità di <i>Authority File</i> per le piattaforme digitali (es. Biblioteche e Archivi digitali) sul web tramite i corrispondenti collegamenti agli <i>Authority Name Files</i> per l'uso normalizzato dei Nomi degli Autori identificati con i valori “non letterali” <i>URIs</i> e pubblicati sul web come entità <i>Linked Open Data</i>
MARC Code List for Relators < http://id.loc.gov/vocabulary/relators.html >	È una lista di termini che descrivono i ruoli di Autori, Relatori, più i loro codici associati per creare le relazioni tra i Nomi propri e le Risorse bibliografiche.
ORCID (Open Researcher and Contributor ID), < http://orcid.org/ >	È un registro avviato nel 2010 come iniziativa no-profit. Consente la registrazione di identificatori univoci (alfa-numeric) per i ricercatori. ORCID risponde al problema, condiviso da molti ricercatori singoli oltre che dalle istituzioni, di collegare in modo attendibile gli autori alle pubblicazioni. ORCID è più di un semplice registro, è la collaborazione di una comunità per integrare questi identificatori nel workflow della ricerca. Prevede due modalità di registrazione degli ID: gli utenti singoli possono registrare e aggiornare il proprio ID gratuitamente, e utilizzarlo in tutte le situazioni in cui possa essere integrato nei workflow tradizionali della ricerca.
FOAF (Friend Of A Friend): < http://xmlns.com/foaf/0.1/ >	È un'ontologia comprensibile dal computer atta a descrivere persone, con le loro attività e le relazioni con altre persone e oggetti. Chiunque può usare FOAF per descriversi. FOAF permette a gruppi di persone di descrivere quel fenomeno noto come <i>Social Network</i> senza la necessità di accedere a un database centralizzato. FOAF è un vocabolario descrittivo espresso tramite il modello dei dati <i>Resource Description Framework</i> (RDF) e <i>Web Ontology Language</i> (OWL). I computer possono usare FOAF, ad esempio, per cercare tutte le persone che vivono in Europa, o

	tutte le persone che hanno un amico in comune, questo appunto perché permette di definire le relazioni tra persone. Ogni profilo ha un identificativo univoco (come ad esempio l'indirizzo email, l'URI dell'homepage o del blog della persona) che viene utilizzato quando vengono definite le relazioni.
“SUBJECT”	
Dewey Decimal Classification (DDC): < http://dewey.info/ >	OCLC (<i>Online Computer Library Center</i>) ha rilasciato il set completo di DDC 23 come <i>Linked Open Data</i> attraverso il servizio <i>Dewey.info</i> , uno spazio online di sperimentazione per DDC in <i>Linked Data</i> (più di 23.000 numeri assegnabili e didascalie in inglese).
Library of Congress (LC) Classification (LCC): < http://id.loc.gov/ >	È un sistema di classificazione bibliografica sviluppato dalla Library of Congress. È utilizzato da gran parte delle biblioteche degli istituti di ricerca e universitari degli Stati Uniti e da molte biblioteche simili in altri paesi, anche se numerose biblioteche pubbliche di dimensioni medio-grandi continuano a servirsi della Classificazione decimale Dewey. Essa essenzialmente è di natura enumerativa, ma suddivide i Soggetti in categorie estese contraddistinguendole con lettere maiuscole. È disponibile anche come <i>Linked data</i> tramite il servizio <i>LC's Linked Data Service</i> (ID.LOC.GOV)
Library of Congress (LC) Subject Headings (LCSH): < http://lcssubjects.org/ >	La versione SKOS di LCSH - disponibile anche come <i>Linked Data</i> tramite il servizio <i>LC's Linked Data Service</i> (ID.LOC.GOV) - è costituita da 30.000 concetti monolingue (inglese) per definire intestazioni di Soggetto. In confronto con Thesauri, intestazioni si riferiscono alle voci pre-coordinate (es. <i>children and terrorism</i>), mentre Thesauri tendono ad utilizzare termini diretti singolari.

<p>The UNESCO Thesaurus: <http://skos.um.es/unescothes></p>	<p>La versione del Thesaurus espressa con i formalismi per il Web SKOS (<i>Simple Knowledge Organization System</i>) è disponibile in quattro lingue (inglese, spagnolo, francese e russo). UNESCO Thesaurus è strutturato in sette settori principali suddivisi in micro-thesauri che consentono all'utente di ottenere una rapida panoramica della materia. Può essere utilizzato per decrivere le risorse nei domini di istruzione, cultura, scienze naturali, scienze umane e sociali, comunicazione e informazione.</p>
<p>Il Thesaurus del Nuovo Soggettario. In formato SKOS/RDF: <http://thes.bncf.firenze.sbn.it/thes-dati.htm></p>	<p>È lo strumento realizzato dalla Biblioteca nazionale centrale di Firenze (BNCF) per l'indicizzazione per soggetto in Biblioteche, Archivi, Mediateche e Centri di documentazione. È costituito da un sistema a più componenti di cui le principali sono: un insieme di Norme sintattiche e semantiche finalizzate all'elaborazione di intestazione di Soggetto; un Thesaurus multidisciplinare in lingua italiana contenente la terminologia impiegabile nell'indicizzazione e, dagli utenti di cataloghi online e altre basedati, nella ricerca per Soggetto.</p> <p>Il Nuovo Soggettario, in quanto strumento di ricerca in rete, si pone nella prospettiva di sviluppare sempre più questo versante, per consentire il collegamento tra il thesaurus e altri tipi di vocabolari ma anche, in senso tecnico, per adottare protocolli Standard e formati adeguati allo scambio di dati, nell'ottica di un'apertura verso domini non librari e nella previsione di integrazione fra Archivi diversi.</p> <p>Per far questo, l'insieme dei metadati del Thesaurus del Nuovo soggettario è stato reso disponibile nel formato SKOS/RDF (mappatura in SKOS v. 0.1 del giugno 2010) ed è accessibile anche attraverso il protocollo Zthes: ciò rientra nell'intenzione della BNCF di promuovere l'uso dei metadati bibliografici anche al di là del tradizionale ambito bibliotecario. Su questo versante, essa ha in corso collaborazioni e contatti a livello nazionale e internazionale per definire le modalità di pubblicazione in rete - come linked data - dei metadati prodotti.</p>
<p>EuroVoc Thesaurus: <http://datahub.io/dataset/eurovoc-in-skos></p>	<p>È un Thesaurus multilingue (23 lingue ufficiali dell'Unione europea) e pluridisciplinare che comprende la terminologia dei settori d'attività</p>

	dell'Unione europea, con un'attenzione particolare per l'attività parlamentare. È passato ad una gestione del Thesaurus basata sull'Ontologia e sulle tecnologie del Web Semantico, in linea con le raccomandazioni del W3C e con gli ultimi sviluppi negli Standard di classificazione.
AGROVOC Thesaurus: < http://datahub.io/dataset/agrovoc-skos >	È un thesaurus strutturato multilingue (22 lingue) che riguarda (con oltre di oltre 30.000 concetti) tutti i campi tematici di agricoltura, silvicoltura, pesca, sicurezza alimentare e settori affini. L'obiettivo principale del dizionario AGROVOC è quello di standardizzare i processi di indicizzazione al fine di facilitare la ricerca all'interno dei sistemi informativi, renderla più efficace e fornire agli utenti i risultati più rilevanti. AGROVOC in linea come <i>Linked Data</i> (essendo anche collegato con altri vocabolari e schemi <i>Linked Data</i>) incoraggia gli sviluppatori dei sistemi di gestione dell'informazione agricola a incorporare AGROVOC nelle loro applicazioni attraverso servizi Web, piuttosto che utilizzare copie locali del database. Con i servizi Web gli aggiornamenti al Thesaurus sono disponibili immediatamente, riducendo il tempo e lo sforzo necessari per scaricare regolarmente ed incorporare l'ultima versione del Thesaurus nell'applicazione.
Library of Congress' Thesaurus for Graphic Materials: < http://id.loc.gov/vocabulary/graphicMaterials.html >	È uno strumento per l'indicizzazione (più di 7.000 termini di soggetto) dei materiali visivi (fotografie, stampe, disegni di progettazione ecc.) Sono aggiunti regolarmente i nuovi termini
STW (Standard-Thesaurus Wirtschaft): Thesaurus for Economics: < http://datahub.io/dataset/stw-thesaurus-for-economics > < http://zbw.eu/beta/sparql/ >	È un Thesaurus bi-lingue (inglese, tedesco), che fornisce il vocabolario su qualsiasi argomento nel campo di economia: circa 6.500 intestazioni per soggetto standardizzati e circa 18.000 termini per definire l'accesso semantico e normalizzato alle risorse.
GEMET (the GEneral Multilingual Environmental Thesaurus): < http://taskman.eionet.europa.eu/projects/zope/wiki/GEMETLinkedData >	È un Thesaurus che copre il dominio di ambiente, ed è disponibile in 29 lingue. La sua versione SKOS è costituita da oltre 5.000 concetti.
Getty Thesaurus of Geographic Names (TGN):	È un vocabolario strutturato di nomi geografici e di altre informazioni per

< http://datahub.io/dataset/getty-tgn >	l'indicizzazione delle risorse nell'arte e nell'architettura.
ONTOLOGIE	
VIVO Ontology: < http://datahub.io/dataset/vivo >	L'Ontologia VIVO è stata finanziata per creare un <i>Facebook semantico</i> per gli scienziati attraverso un'Ontologia di base (per la modellazione di scienziati, pubblicazioni, risorse, grant, località e servizi). La versione 1.0 di VIVO contiene 236 classi con 278 proprietà e 222 proprietà dei dati, incorporando anche le Classi da ontologie come BIBO, Dublin Core, FOAF e SKOS. I dati VIVO sono focalizzati sulla ricerca (es. titoli di studio, pubblicazioni, esperienza, borse di studio), sulla didattica (es. corsi, seminari, corsi di formazione), sui servizi (es. l'organizzazione di conferenze, tavole editoriali, altri servizi comunitari).
Bibliographic Ontology (BIBO): < http://bibliontology.com/specification > < http://purl.org/ontology/bibo/ > <i>- puo' essere usata come una ontologia citazionale o per la classificazione documentale per descrivere qualsiasi tipo di risorse bibliografiche in RDF -</i>	Questa Ontologia fornisce i concetti e le proprietà principali per la descrizione di citazioni e riferimenti bibliografici (es. citazioni, libri, articoli) sul Web Semantico. Altri vocabolari (es. elementi di metadati DC per una semplice descrizione bibliografica, FOAF) possono essere combinati con i termini BIBO, come estensioni locali.
GeoNames Semantic Web: < http://datahub.io/it/dataset/geonames-semantic-web >	L'Ontologia GeoNames permette di aggiungere informazioni semantiche geospaziali alle risorse web. Contiene oltre 6,2 milioni di toponimi GeoNames che hanno un identificatore di localizzazione URL univoco con un servizio web RDF corrispondente. Descrive le relazione tra toponimi.
Linee Guida OpenAIRE (V 2.0) raccomandano usare i valori dagli schemi di Classificazione SURF: < http://wiki.surf.nl/display/standards/info-eu-repo/#info-eu-repo-ClassificationSchemes/ >	
Namespace/il dominio del Vocabolario sul Web	Vocabolario come schema di codifica
< info:eu-repo/classification/ddc >	Dewey Decimal Classification (DDC): < http://www.oclc.org/dewey/ >
< info:eu-repo/classification/udc >	Universal Decimal Classification: < http://www.udcc.org/http://www.udcc.org/ >
< info:eu-repo/classification/lcc >	Library of Congress Classification: < http://www.loc.gov/catdir/cpsolcco/http://www.loc.gov/catdir/cpsolcco/ >

	>
<info:eu-repo/classification/lcsh>	Library of Congress Subject Headings (LCSH)
<info:eu-repo/classification/swd>	Schlagwortnormdatei (SWD)
<info:eu-repo/classification/biessc>	Book Industry Communication - Standard Subject Categories: < http://www.bic.org.uk/7/BIC-Standard-Subject-Categories/ >
<info:eu-repo/classification/nbc>	Nederlandse Basisclassificatie (Dutch Basic Classification): < http://www.kb.nl/vak/basis/bc04.pdf >
<info:eu-repo/classification/jel>	Sistema di Classificazione <i>Journal of Economic Literature</i> (JEL) Gli Articoli nelle Riviste delle scienze economiche di solito vengono classificati secondo i codici della Classificazione JEL:< http://www.aeaweb.org/journal/jel_class_system.php >

Tabella 14. Alcuni Vocabolari controllati e Schemi semantici per la normalizzazione e l'abilitazione dei metadati d'Archivio come *Linked Data*

5.3. Alcuni approcci, documenti e principi per la creazione di metadati estensibili di qualità per descrivere i contenuti di ricerca

Nelle varie fasi del ciclo di vita di un oggetto informativo, i suoi autori, gestori e utilizzatori dovrebbero creare intorno alla descrizione del medesimo oggetto più metadati possibile. Questi agenti dovrebbe essere anche consapevoli che, a seconda della natura delle risorse e delle loro collezioni, un unico schema di metadati non può mai essere sufficiente per tutte le esigenze della loro descrizione. Una logica e corretta combinazione di elementi di metadati provenienti da diversi schemi può essere la soluzione migliore per dare un giusto supporto alla rappresentazione, gestione, ricerca e conservazione dei materiali digitali di varia tipologia [V. sito: *NISO Framework Working Group*].

Lo schema dei metadati proposto dalle “Linee Guida per la creazione e la gestione di metadati nei Repository Istituzionali” (CRUI) [V. sito] offre l’utilizzo esteso dei 15 metadati dello schema *Dublin Core* (DC) semplice [V. APPENDICE A, § 1.1.3]. Il DC semplice presenta gli elementi di metadati di base per descrivere oggetti di contenuto (risorse informative, *Content Objects*) negli Archivi Istituzionali, al fine di supportare l’interoperabilità minima tra i Repository OAI (*Open Archive Initiative*) - compatibili attraverso il protocollo OAI-PMH [V. PARTE I, Cap. 2, § 2.1].

I fattori che spingono diversi agenti sul Web a utilizzare i metadati DC sono:

- la semplicità di questi metadati, per cui quasi chiunque può usarli (o almeno parti di essi);
- DC è una scelta per *default* per la descrizione delle risorse in diversi Archivi istituzionali;
- l’alta capacità di integrazione del DC con altri Profili di metadati (es. *DC-Library Application Profile*, *Scholarly Works Application Profile*, *VOA3R AgRes AP Metadata Terms*).

Una volta che metadati DC sono estesi tramite gli opportuni qualificatori, i loro valori possono essere comunque appiattiti (impoveriti) dai meccanismi di raccolta dei metadati entro le piattaforme OAI tramite il protocollo OAI-PMH. Per contrastare tali effetti, sia *Data Provider* (Archivi) che *Service Provider* (Archivi aggregatori di dati da *Data Provider*) dovrebbero sostenere Standard e Protocolli comuni e largamente condivisi, così come investire nello sviluppo degli Schemi di metadati incrociati tramite le mappature tra loro elementi, limitando la perdita dei metadati e della specificità dei loro valori.

È una buona pratica quando negli Archivi Istituzionali la scelta motivata degli elementi dei metadati, insieme con la compilazione coerente dei loro valori, approcci di progettazione, requisiti di codifica e conservazione vengono dichiarati nelle apposite Politiche. Queste ultime sono anche molto importanti per lo sviluppo di un nuovo metodo a supporto del ciclo di vita dei contenuti digitali (incluso Dati di ricerca) [V. sito: *Research Data*] – quale *Data Management Plan* (DMP) [V. sito].

Insieme ai metadati di base qualitativamente programmati, codificati e largamente incrociati (mappati) con altri schemi, “*Institutional Repositories will be ultimately to form an International Network of Indexed Repositories searchable from a Single Interface*” [V. sito: *Statement from the University of Oregon Libraries*].

Durante la selezione e lo sviluppo dei valori di metadati sarebbe opportuno fare un confronto continuo con sei principi formulati da NISO (*National Information Standards Organization*) [V. sito: *NISO Framework Working Group*] che riguardano la creazione di “*Good metadata*”, ossia metadati con requisiti di qualità. Occorre anche riportare che nel 2013 NISO ha lanciato un’iniziativa per sviluppare Standard per i “*Metadati e Indicatori Open Access*”, particolarmente per i dati bibliografici e gli identificatori alla descrizione dell’accessibilità agli articoli di Riviste nel rispetto delle condizioni della loro apertura (*Open Access*) [V. sito: *NISO Launches New Initiative*].

Secondo NISO, "Good metadata":

1. sono conformi agli Standard comunitari per essere appropriati e messi in atto adeguatamente per la descrizione delle risorse e delle loro collezioni, e per soddisfare le esigenze informative degli utenti (attuali, futuri e potenziali);
2. supportano l'interoperabilità tra le risorse che descrivono;
3. usano gli strumenti del controllo di Autorità e Standard per collocare oggetti correlati;
4. includono chiare dichiarazioni sulle condizioni e termini d'uso per ogni oggetto digitale;
5. sostengono conservazione di oggetti in collezioni a lungo termine.
6. sono anche essi oggetti che devono avere i requisiti di oggetti di qualità, tra cui autorità, autenticità, archiviazione, persistenza e identificazione univoca.

"Good metadata" richiedono la comprensione dei dati che stanno per descrivere e Standard con i quali le medesime descrizioni sarebbero state possibili.

Nella sezione "Validazione di metadati" delle "Linee Guida per la creazione e la gestione di metadati nei Repository Istituzionali" (CRUI) [V. sito, pp. 11-12] viene sottolineato che la qualità dei metadati, a sua volta, determina la qualità delle funzioni svolte e dei servizi offerti sia dai Repository (*Data Provider*, Archivi come fornitori di dati) che dagli aggregatori (*Service Provider*) che raccolgono i dati dai Data Provider via il Protocollo dell'interoperabilità OAI-PMH, considerando il contesto delle piattaforme conformi a *Open Archive Initiative*.

Nella creazione di "Good metadata", è opportuno anche riferirsi ai documenti di autorità come "User Guide/Creating Metadata" [V. sito] sviluppato all'interno della Comunità Dublin Core Metadata Initiative (DCMI) e "Meaningful Bibliographic Metadata" (M2B) [Subirats, Zeng, 2012].

5.3.1.Strategie delle "Linee guida per la creazione e la gestione di metadati nei Repository Istituzionali" (CRUI, 2012)

Per garantire l'accuratezza dei metadati e la loro compilazione qualitativa durante il processo di auto-archiviazione dei materiali digitali negli Archivi Istituzionali, le Linee guida CRUI raccomandano di:

1. Assistere gli utenti durante l'auto-archiviazione (basata sul processo di inserimento di metadati descrittivi e di caricamento di file) degli oggetti negli Archivi Istituzionali. Questa assistenza può essere espressa tramite la creazione di strutture come *Metadata Editor* con le liste dinamiche di auto-completamento e l'importazione dei valori di metadati provenienti da diverse fonti autorevoli (es. per qualificare i valori di metadati "Responsible Body", "Subject", "Place").
2. Validare metadati inseriti prima della loro esposizione agli utenti finali e ai fornitori dei servizi. L'efficacia e l'efficienza dei metadati in relazione alla loro importazione ed esportazione dall'Archivio sono strettamente correlate all'uso di *identificatori univoci e persistenti* come DOI, HANDLE basati sugli indirizzi Web URIs, per assicurare la stabilità di elementi di metadati, così come la loro interoperabilità entro e tra sistemi diversi accessibili tramite il Web.

Con la validazione dei metadati, la duplicazione dei loro valori sarà facilmente evitata, e possono essere sviluppati efficaci filtri per la navigazione e scoperta di risorse correlate (es. tramite i valori connessi allo stesso "Responsible Body", "Subject", "Place").

Gli identificatori potrebbero essere anche di grande importanza nella creazione delle connessioni tra contenuti di ricerca e loro processi di valutazione. In questo modo gli Archivi Istituzionali possono fungere anche da infrastrutture tecniche per la gestione e valutazione delle attività di ricerca [V. sito: *Institutional Repositories for Research Management and Assessment*].

Sfruttando con l'efficacia e l'efficienza all'interno delle reti Web il potenziale:

- degli *identificatori univoci e persistenti* assegnati alle risorse di Archivi Istituzionali,
- dello standard CERIF (*Common Information Format europeo della ricerca*) e
- di altri Standard, modelli e Profili Applicativi (es. SWAP: *Scholarly Work Application Profile*) di metadati a supporto dei contenuti di ricerca,
- di Vocabolari controllati e Classificazioni comprese anche quelle scientifico-disciplinari - può essere raggiunta una maggiore integrazione tra gli Archivi Istituzionali e i Sistemi Informativi di ricerca CRIS e Euro-CRIS. Per supportare questa integrazione, nel 2013 la Confederazione internazionale COAR (*Confederation of Open Access Repositories*) e i Sistemi Informativi di ricerca europei Euro-CRIS hanno annunciato il loro strategico partenariato [V. sito: *Euro-CRIS and COAR join forces*].

Attualmente, diversi sistemi informativi di ricerca (*Research Information Systems, RIS*) a cui si collega una varietà di Archivi Istituzionali implementano assieme il succitato modello dei dati CERIF che esprime concetti e relazioni di collegamento (con il ruolo e la durata temporale) di diverse entità coinvolte nel processo di ricerca scientifica.

Diversi RIS pubblicano API Web utilizzando tecnologie SOAP o REST, facendo leva sull'importanza dell'uso delle tecnologie come RDF, SPARQL, ontologie Web come OWL, BIBO, FOAF, per supportare le applicazioni web e *mash-up* (fusione) con i dati provenienti da altri sistemi, incluso gli Archivi Istituzionali [Jeffery, Corson-Rikert, 2012, p.17].

3. Fornire ogni Repository con il supporto professionale in quanto allo sviluppo e alla compilazione qualitativi di metadati.

Nel prossimo futuro è molto probabile che tutti i Repository locali (*Data Providers*, qui Archivi Istituzionali) saranno costretti ad impiegare nelle loro piattaforme una descrizione qualitativa dei metadati in quanto i database di citazioni più importanti e gli aggregatori (*Service Providers*) dei contenuti di ricerca devono realizzare le loro missioni di citazione, aggregazione e distribuzione di dati tramite i servizi a valore aggiunto secondo gli Standard di alta qualità [Šimek, p. 88]

Il concetto di "metadati di qualità" richiama il concetto di un ambiente "*Trusted*" ovvero affidabile che viene attivamente promosso nel quadro di *Trusted Digital Repositories* [V. sito] ovvero gli Archivi sviluppati in relazione ai requisiti degli Standard, delle Raccomandazioni e Linee ampiamente riconosciuti, condivisi e impiegati da diversi soggetti produttori dei contenuti informativi.

Gli elementi di metadati presentati dalle Linee Guida CRUI mirano a coprire una descrizione di base dei seguenti tipi di risorse digitali di ricerca: Articolo, Patente, Libro e parte del libro, Articolo di Conferenza, Poster di Conferenza, Paper di convegno, Annotazione, Recensione, Tesi di Dottorato, Tesi di Laurea, *Working Paper*.

Lo schema dei metadati presentato nel penultimo paragrafo del presente Capitolo introdurrà:

- 1) gli elementi di metadati dalle "Linee Guida per la creazione e la gestione di metadati nei Repository Istituzionali" (CRUI, 2012) raggruppati secondo le categorie di metadati con le proprietà comuni definiti dalle Raccomandazioni "LODE-BD" (*Linked Open Data – enabled Bibliographical Data*) (2012);
- 2) un numero esteso di elementi di metadati definiti dalle Linee Guida CRUI, in conformità con le mappature tra gli elementi di metadati proposte dalle Raccomandazioni "LODE-BD";
- 3) per ogni elemento di metadati - le raccomandazioni di compilazioni e codifica secondo le strategie "LODE-BD".

In Tabella 15 vengono brevemente presentati alcuni autorevoli documenti considerati dalle "Linee Guida per la creazione e la gestione di metadati nei *Repository* Istituzionali" durante il processo della creazione dello schema di metadati presentato dalle medesime Linee.

1.	Dublin Core (DC) -Library Application Profile (DC-Lib): < http://dublincore.org/documents/library-application-profile/ >
2.	Using Dublin Core - Dublin Core Qualifiers: < http://dublincore.org/documents/2005/11/07/usageguide/qualifiers.shtml >
3.	DRIVER Guidelines v.2.0 sviluppati da SURF (un'organizzazione Olandese non-profit per l'innovazione ICT in ambito universitario e di ricerca: < http://www.surf.nl/en/oversurf/Pages/Default.aspx > L'uso di mappature tra le tipologie documentali: < http://www.driver-support.eu/documents/DRIVER_Guidelines_v2_Final_2008-11-13.pdf >
4.	SWAP: Scholarly Works Application Profile: < http://www.ukoln.ac.uk/repositories/digirep/index/Scholarly_Works_Application_Profile > La mappatura degli elementi dei metadati proposti dalle “Linee guida per la creazione e la gestione di metadati nei Repository Istituzionali” (CRUI) con gli elementi SWAP è raggiunta tramite la mappatura prevista da SURF
5.	OpenAIRE Guidelines (Versione 1.1): < http://www.openaire.eu/en/component/content/article/207 > Queste Linee guida sono costruiti sulle Linee guida DRIVER. Forniscono le strategie per consentire agli Archivi ad Accesso Aperto di esporre in modo più coerente i campi di metadati comuni circa i prodotti di ricerca finanziati, finanziamenti e livelli di accesso del materiale di ricerca. La Versione 1.1 è aggiornata con la Versione 2.0: < http://www.openaire.eu/en/component/content/article/427-openaire-releases-version-20-of-the-openaire-guidelines >
6.	PRISM 3.0 Draft Metadata Specifications: < http://www.idealliance.org/downloads/prism-30-draft-metadata-specs > Queste specifiche definiscono i campi di metadati strutturati XML e i Vocabolari controllati per la gestione, l'aggregazione e distribuzione dei contenuti pubblicati nelle Riviste
7.	<info:eu-repo> namespaces: < http://wiki.surffoundation.nl/display/standards/info-eu-repo > Questo è il namespace di identificatori utilizzati dall' <i>European Repository Systems</i> , per descrivere le relazioni semantiche in documenti composti e complessi
8.	ISO 214 Standard (Documentation. ABSTRACTS for publications and documentation): < http://www.univr.it/documenti/Documento/allegati/allegati112794.pdf >

9.	<p>International Standards to CITING publications:</p> <ul style="list-style-type: none"> - APA (<i>American Psychological Association</i>): <http://www.indiana.edu/~citing/APA.pdf> - Chicago (<i>Chicago Citation Style</i>): <http://www.libraries.iub.edu/secure/defiles/Chicago.pdf> - MLA (<i>Modern Language Association</i>): <http://www.indiana.edu/~citing/MLA.pdf>
10.	<p>ISO 639-1, ISO 639-2, ISO 639-3 Standards (Codes for the representation of names of Languages): <http://www.loc.gov/standards/iso639-2/php/code_list.php></p>
11.	<p>Schemi di autorità prodotti nel contesto italiano:</p> <ul style="list-style-type: none"> - schema REICAT (nuove regole di catalogazione italiana): <http://www.iccu.sbn.it/upload/documenti/REICA_bozza_complessiva_genn2009.pdf> - Il Thesaurus del Nuovo Soggettario: <http://thes.bncf.firenze.sbn.it/ricerca.php> - Classificazione dei Settori Scientifico-Disciplinari (SSD): <http://www.cun.it/media/116411/settori_scientifico_disciplinari_english.pdf> - Schema ISI-CRUI per aree disciplinari: <http://www.univr.it/documenti/Documento/allegati/allegati305644.pdf>; <http://www.univr.it/documenti/Documento/allegati/allegati112794.pdf>

Tabella 15. Alcuni autorevoli documenti considerati dalle “Linee Guida per la creazione e la gestione di metadati nei Repository Istituzionali” (CRUI, 2012)

5.3.2. Strategie delle Raccomandazioni “LODE-BD” (*Linked Open Data – enabled Bibliographical Data*)

Le Raccomandazioni “LODE-BD” comprendono importanti suggerimenti per un fornitore di dati (*Data Provider*, Archivio) che deve produrre i metadati strutturati, abilitati e condivisibili come *Linked Open Data* (LOD). Questi metadati possono descrivere le risorse bibliografiche come Articoli, Monografie, Tesi, testi di Conferenze, materiali di presentazione, rapporti di ricerca, materiale didattico [V. sito: *Linked Open Data – enabled Bibliographical Data*, p. 4].

Secondo “LODE-BD”, per la descrizione delle risorse di ricerca in un Archivio digitale:

1) i metadati Dublin Core dai namespace *DCMES* (*Dublin Core Metadata Element Set*) (DC:) e *DCMI* (*Dublin Core® Metadata Initiative*) *Metadata Terms* (DCTERMS:) sono fondamentali, mentre i termini di metadati provenienti da altri namespace sono supplementari, utili per soddisfare eventuali ulteriori esigenze per la descrizione delle risorse del Repository.

2) Per estendere l’espressività dei metadati di base, essi possono essere mappati con gli elementi di BIBO Ontology, AGLS *Metadata Standard of the Australian Government Locator Service*, eprint (UKOLN *Eprints Terms*, SWAP), MARCrel (*MARC List for Relators*).

In vista di queste considerazioni, “LODE-BD” affrontano due principali questioni:

1. il modo in cui i dati - ospitati da diversi Archivi aperti (*Data* e *Service Provider*) - possono essere meglio scambiati tra loro;
2. la modalità con la quale codificare questi dati abilitandoli come *Linked Open Data* (*LOD-enabled*). “LODE-BD” fornisce un numero selezionato di Standard di metadati e di Vocabolari *LOD* (*LOD-enabled*) con cui poter affiancare la codifica dei valori di metadati.

Tutti gli elementi dei metadati sono presentati in una Tabella che offre gli spazi per le mappature dei metadati descrittivi DC con i metadati degli schemi supplementari e anche con i valori dei vocabolari controllati, con cui saranno ovviamente raggiunti i benefici in termini di coerenza, estensibilità semantica e ricchezza dei valori di metadati di base.

Facendo riferimento alla fase dello sviluppo dei valori di metadati secondo le Raccomandazioni “LODE-BD”, i gestori di *Repository* dovrebbero affrontare le seguenti questioni:

1) Quali tipi di entità e relazioni devono essere coinvolte nel descrivere e accedere le risorse bibliografiche?

2) Quali caratteristiche a livello locale dovrebbero essere considerate per la pubblicazione di dati bibliografici *LOD-ready* ovvero di metadata abilitati come *LOD*?

3) Quali sono i valori di metadati appropriati in ogni data proprietà, quando la produzione dei metadati richiede la loro abilitazione come *LOD*? [V. sito: *Linked Open Data – enabled Bibliographical Data*, p. 5].

5.3.2.1. Alberi decisionali “*Decisions Trees*” : scelta tra valori ”letterali” e ”non-letterali” per la codifica delle proprietà di metadati

La vera forza dei passi proposti dalle Raccomandazioni “LODE-DB” per lo sviluppo dei metadata sono gli “Alberi decisionali” o “*Decision Trees*” (Figura 54) [V. sito: *Linked Open Data – enabled Bibliographical Data*, p. 31-32] designati per agevolare la selezione delle strategie appropriate da parte di *Data Provider* (fornitori di dati) in base alle loro esigenze locali, con l’obiettivo di raggiungere un alto grado di scambio e di riuso dei valori di metadati nel Web di Dati, ovvero nello spazio *Linked (Open) Data* – espressione pratica del Web Semantico.

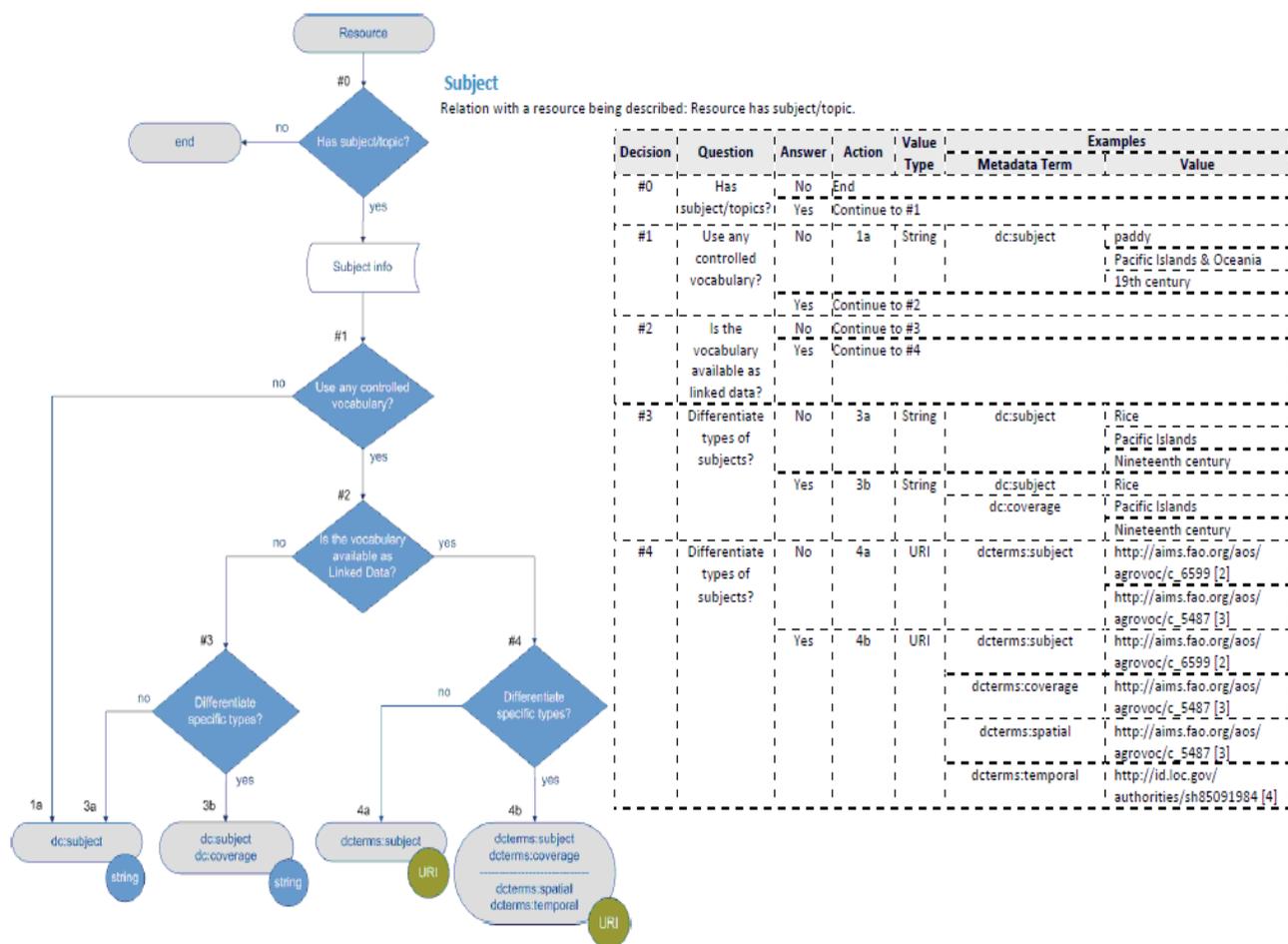


Figura 54. Albero decisionale "LODE-BD" e tabella esplicativa a supporto dei passi della codifica dei valori del metadata "Subject"

Gli Alberi decisionali "LODE-BD" sono sviluppati per appoggiare la codifica e la descrizione dei seguenti elementi di metadata suddivisi in nove gruppi:

- | | |
|---------------------------------|---|
| 1. Title information | Title/Alternative title |
| 2. Responsible Body | Creator. Contributor. Publisher |
| 3. Physical Characteristics | Date. Identifier. Language. Format/Medium.
Edition/Version. Source |
| 4. Holding/Location information | Location/Availability |
| 5. Subject Information | Subject/Topic |
| 6. Description of Content | Description/Abstract/Table of Contents
Type/Form/Genre |
| 7. Intellectual property rights | Right Statements |
| 8. Usage | Audience/literary indication/ education level |
| 9. Relations | Relation between resources. Relation between agents |

Tutti gli Alberi decisionali "LODE-BD" iniziano dalla proprietà che descrive l'istanza "Risorsa" e vengono distribuiti in diagrammi di flusso che offrono la scelta dei vari passi per la codifica dei metadata. Alla fine di ogni Albero decisionale viene fornita una Tabella di testo contenente note, passi ed esempi in corrispondenza ai suggerimenti per i passi di codifica, dove è essenziale. La

codifica delle proprietà di metadati può avvenire tramite due tipi di valori: "letterale" e "non letterale" (Figura 50).

1) **Valore letterale.** Questo valore è in genere una stringa di caratteri che utilizza come forma lessicale una stringa UNICODE, insieme a un <tag> opzionale per indicare lingua o tipo di dati che denotano una "Risorsa".

Esempi di *namespace* di metadati con Valori *letterali*

dcterms:alternative "A Feast of Beans"
dcterms:available"2006-07"^^dcterms:W3CDTF ...

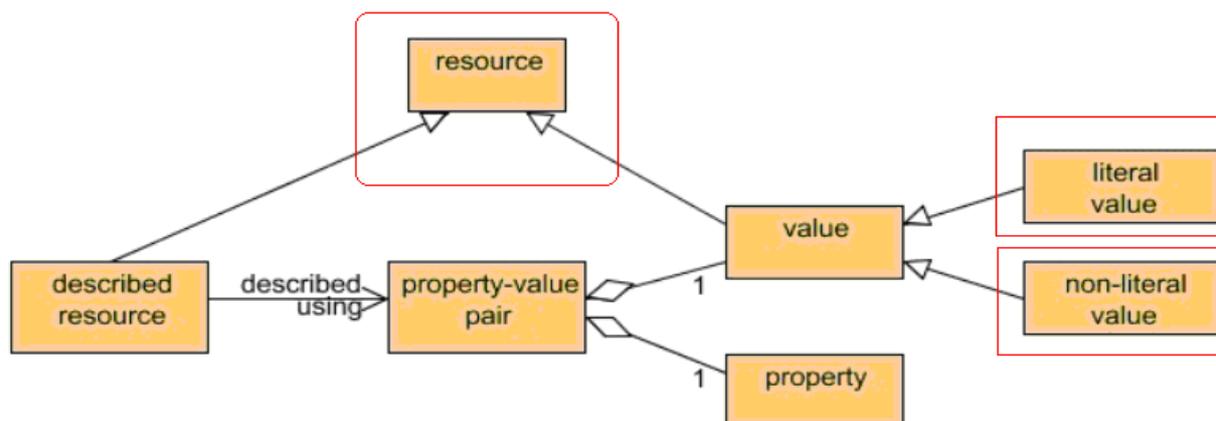


Figura 55. Valori: "letterali" e "non letterali" per qualificare le proprietà di metadati descrittivi risorse informative

2) **Valore non letterale.** Questo valore presenta entità fisiche, digitali, o concettuali indicati dagli Identificatori univoci.

Esempi di *namespace* di metadati con Valori *non letterali*

dcterms:conformsTo
 <<http://www.w3.org/2001/XMLSchema>>
dcterms:contributor gnd:135066719
 gnd:135066719
 foaf:familyName "Elliott";
 foaf:givenName "Missy" ;
 foaf:nick "Missy E"...

Gli "Alberi decisionali"LODE-BD aiutano fornitori di dati (*Data Provider*) di valutare il divario esistente tra l'uso corrente di "valori letterali" e la loro evoluzione verso un approccio *Linked Data* facendo l'uso dei valori "non letterali" basati sugli indirizzi Web *http_URI* in riferimento ai vocabolari controllati, alle Ontologie e ad altri set con i valori "non letterali".

Proprietà di alcuni *namespace* di metadati DC ("DC:" e "DCTERMS:") - come è dimostrato nelle Tabelle esplicative "LODE-BD" - possono essere qualificati sia con i valori "letterali" che con quelli "non-letterali".

Tuttavia, per la produzione dei metadati abilitati come *Linked Open Data* da parte di *Data Provider*, affinché i medesimi metadati possano essere facilmente raccolti dagli aggregatori e dai fornitori di servizi in rete *Service Provider*, è raccomandato ricorrere all'uso di metadati sotto il *namespace* "DCTERMS:" che può accogliere la codifica di valori "non-letterali" (URIs) [V. sito: *Dublin Core User Guide/Publishing Metadata as Linked Data.*].

La rilevanza pragmatica dell'approccio degli "Alberi decisionali"LODE-BD per la produzione dei metadati abilitati come *Linked Open Data* è che ogni fornitore di dati (*Data Provider*) può evidenziare all'interno di un concreto "Albero decisionale" i propri percorsi decisionali, scegliendo gli elementi di metadati da utilizzare così come i vocabolari e le norme a loro sostegno.

Gli Alberi decisionali "LODE-BD" non sono limitati ai domini specifici, e possono essere appropriati per l'utilizzo da qualsiasi fornitore di dati in conformità alle esigenze locali. Tuttavia, le decisioni riguardanti quali Standard, Vocabolari e Schemi adottare avranno un impatto diretto sulla prontezza dei metadati a diventare *Linked Open Data*, ovvero ad avere il requisito di "LOD-readiness".

5.3.2.2. Alcuni vocabolari controllati per l'abilitazione di metadati *RightsHolder*, *Licence* come *Linked Data*

Prima che una certa risorsa venga pubblicata, è importante decidere sotto quale Licenza essa sarà presentata agli utenti. Come è stato già menzionato in connessione allo Schema che riguarda la produzione *Linked Open Data* - "Tim's 5 star deployment scheme", per contribuire ad una migliore visibilità, usabilità e interoperabilità dei contenuti sul Web, si consiglia di pubblicarli sotto Licenze Aperte (*Open License*), tramite cui i contenuti possono essere liberamente:

- condivisi (copiati, distribuiti e trasmessi);
- combinati;
- utilizzati da qualsiasi terza parte (anche commerciale) per la produzione dei derivati, in ogni caso con l'attribuzione dell'opera all'Autore o al Licenziante.

Di buona norma, alle risorse di ricerca prodotte con i fondi pubblici devono essere applicate le Licenze Aperte [De Robbio, 2012]. Tuttavia, si deve prendere anche in considerazione che alcune risorse di Archivi Istituzionali possono essere collegati con le questioni di:

- "Accesso di Embargo"/"Embargoed Access" (la risorsa è di Accesso Chiuso, fino che venga rilasciata come *Open Access* in una certa data);
- "Accesso Limitato"/"Restricted access" (*Open Access*, ma con le restrizioni) e
- "Accesso Chiuso" (opposto all'*Open Access*).

Quindi, per denotare le proprietà dei metadati "*dcterms:rightsHolder*", "*dcterms:licence*" - possono essere usati sia i riferimenti ai valori di *Licenze Aperte* che chiuse.

Gli autori che vogliono pubblicare i loro lavori di ricerca e successivamente depositarli in un Archivio Aperto *Open Access*, possono trovare tutte le informazioni utili a supporto informativo circa i "Diritti di proprietà intellettuale" e le "Licenze" all'interno dei servizi online come:

- 1) "*SHERPA/Romeo*" [V. sito] che, a partire dal 2006, raccoglie in un database tutte le Politiche adottate dai diversi Editori nei confronti dell'auto-archiviazione degli articoli pubblicate nelle loro riviste;
- 2) "*Diritto d'autore*" offerto dal Sistema Bibliotecario dell'Università degli Studi di Padova [V. sito];
- 3) "*Diritto d'autore*", offerto da PLEADI (Portale per la Letteratura scientifica Elettronica Italiana su Archivi aperti e Depositi Istituzionali) [V. sito].

Alcuni passi decisionali su come scegliere una certa Licenza per descrivere l'uso delle Risorse in un Archivio Istituzionale sono presentati in Figura 56. I medesimi passi hanno affiancato l'Albero Decisionale LODE-BD riferito al metadato "Diritti" che offre una guida nelle situazioni e buone pratiche per la codifica dei dati riferiti a questo Metadato.

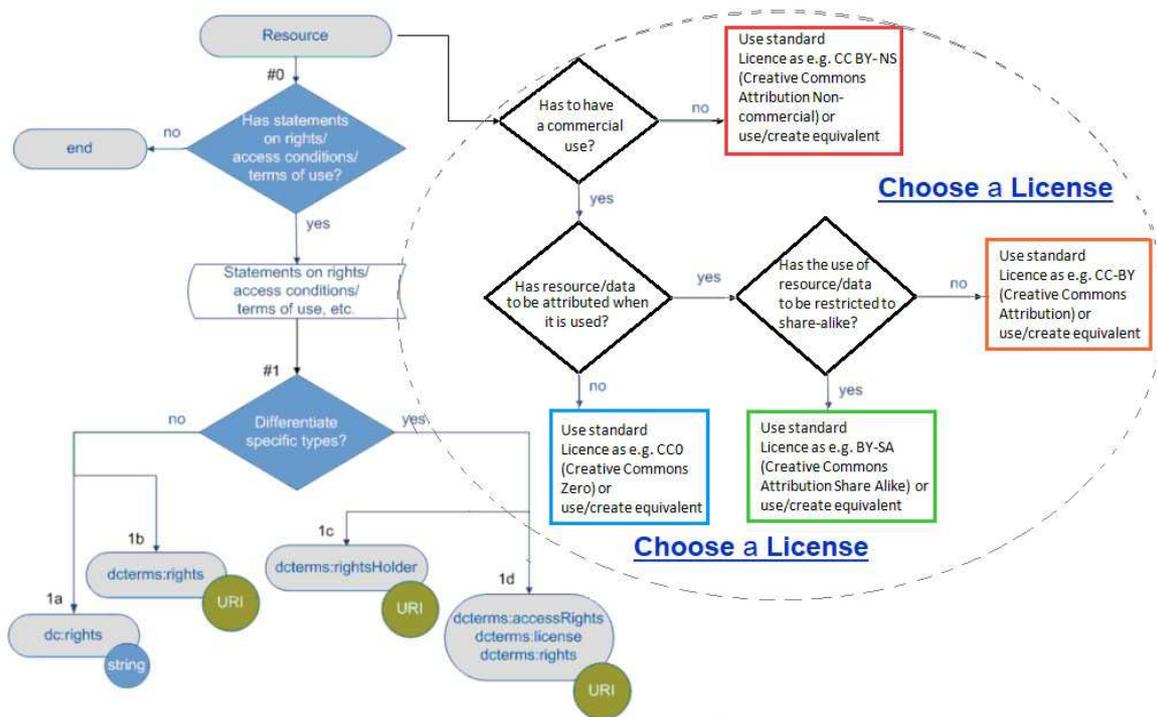


Figura 56. Passi decisionali per la scelta di Licenze a supporto di pubblicazione di diverse risorse

L'estensione fornita all'Albero Decisionale presentato in Figura 56 può essere interpretata come segue:

<p>La Risorsa che deve essere depositata nell'Archivio prevede un uso Commerciale?</p>	<p>No</p> <p>↑</p> <p>Si</p>	<p>Utilizzare Licenza Standard (es. CC BY-NS) o crearne una specifica (per soddisfare le esigenze locali), compatibile con lo Standard</p>
<p>Alla Risorsa dell'Archivio di proprietà intellettuale istituzionale deve essere attribuito un uso aperto?</p>	<p>No</p> <p>↑</p> <p>Si</p>	<p>Utilizzare Licenza Standard (es. CC0, ODC PDDL) o creare una specifica compatibile con lo Standard</p>
<p>Alla Risorsa deve essere attribuito il requisito "share-alike"?</p>	<p>No</p> <p>↑</p> <p>Si</p>	<p>Utilizzare Licenza Standard (es. CC-BY) o creare una specifica compatibile con lo Standard</p> <p>Utilizzare Licenza Standard (es. BY-SA) o creare una specifica compatibile con lo Standard</p>

Dopo che una certa Licenza è stata scelta, un valore ("letterale" e / o "non letterale") che identifica formalmente il tipo della Licenza deve essere codificato nella proprietà del termine di Metadato "Licence", come secondo le strategie di codifica "LODE-BD" [V. sito: "LODE-BD": 3.7. Rights].

In Tabella 16 vengono presentati alcune Licenze "Open" e "Not-open" che possono essere considerate come Vocabolari controllati identificati con gli indirizzi Web http_URIs che possono abilitare i metadati "RightsHolder", "Licence" come *Linked Data*.

Codifica delle proprietà dei metadati “ <i>dcterms:rightsHolder</i> ”, “ <i>dcterms:license</i> ”		
Vocabolari che identificano le Licenze con URIs. Validi per l’abilitazione dei metadata come <i>Linked Data</i>	Tipologia	Identificazione con URIs
Creative Commons Attribution (CC BY)	<i>Open</i>	< http://creativecommons.org/licenses/by/3.0/legalcode > <i>Descrizione dettagliata:</i> < http://creativecommons.org/licenses/by/3.0/ >
Creative Commons Attribution Share Alike (CC BY-SA)	<i>Open</i>	< http://creativecommons.org/licenses/by-sa/3.0/ > <i>Descrizione dettagliata:</i> < http://creativecommons.org/licenses/by-sa/3.0/legalcode >
Open Data Commons Public Domain Dedication and License (ODC PDDL)	<i>Open</i>	< http://opendatacommons.org/licenses/pddl/1-0/ >
Creative Commons CC Zero (CC0)	<i>Open</i>	< http://creativecommons.org/publicdomain/zero/1.0/legalcode > <i>Descrizione dettagliata:</i> < http://creativecommons.org/publicdomain/zero/1.0/ >
GNU Free Documentation License (GNU FDL)	<i>Open</i>	< http://www.gnu.org/copyleft/fdl.html >
Open Database License (ODbL)	<i>Open</i>	< http://opendatacommons.org/licenses/odbl/1.0/ >
GNU General Public License (GNU GPL)	<i>Open</i>	< http://www.gnu.org/copyleft/gpl.html >
New BSD License	<i>Open</i>	< http://en.wikipedia.org/wiki/BSD_licenses#3-clause_license_.28.22New_BSD_License.22_or_.22Modified_BSD_License.22.29 >
Simplified BSD License	<i>Open</i>	< http://en.wikipedia.org/wiki/BSD_licenses#2-clause_license_.28.22Simplified_BSD_License.22_or_.22FreeBSD_License.22.29 >

Creative Commons Attribution Non-commercial (CC BY-NC)	<i>Not open</i>	< http://creativecommons.org/licenses/by-nc/3.0/legalcode > <i>Descrizione dettagliata:</i> < http://creativecommons.org/licenses/by-nc/3.0/ >
Per approfondire la questione dell'impiego di diverse Licenze V. De Robbio [2011 a].		
Codifica delle proprietà di metadati “<i>dcterms:accessRights</i>”, “<i>dcterms:rights</i>” in conformità con le Linee Guida <i>OpenAIRE</i> (V 2.0)		
Vocabolari che identificano le Tipologie di accesso con URIs. Validi per l'abilitazione dei metadata come <i>Linked Data</i>	Tipologia	Descrizione
< info:eu-repo/semantics/openAccess >	<i>openAccess</i>	< http://wiki.surf.nl/display/standards/info-eu-repo/#info-eu-repo-AccessRights >
< info:eu-repo/semantics/restrictedAccess >	<i>restricted Access</i>	
< info:eu-repo/semantics/closedAccess >	<i>closedAccess</i>	
< info:eu-repo/semantics/embargoedAccess >	<i>embargoed Access</i>	
Le Linee Guida <i>OpenAIRE</i> (V 2.0) Raccomandano l'uso dei valori dal vocabolario per i Diritti d'accesso “ <i>SURF vocabulary</i> ” in riferimento al <i>namespace</i> < http://wiki.surf.nl/display/standards/info-eu-repo/ > Secondo <i>OpenAIRE Guidelines</i> V 2.0, quando “ <i>accessRights</i> ” = “ <i>embargoEnd</i> ”, è raccomandato usare le date riferite alla fine del periodo di embargo (<i>embargo end date</i>)		
< info:eu-repo/date/embargoEnd >	<i>embargoEnd</i>	< http://wiki.surf.nl/display/standards/info-eu-repo/#info-eu-repo-DateTypesandvalue >

Tabella 16. Alcuni vocabolari controllati per l'abilitazione dei metadati “*RightsHolder*”, “*Licence*”, “*AccessRights*”, “*Rights*” come *Linked Data*

5.3.2.3. Mappatura dei metadati con i marcatori "Schema.Org" per una migliore comunicazione dei contenuti sul Web

Nel'Appendice 4 delle Raccomandazioni "LODE-BD" (*Linked Open Data-enabled Bibliographical Data Recommendations*) [V.sito] vengono incrociati (mappati) gli elementi di alcuni metadati con i marcatori (*mark-up*) dello "Schema.org" [V.sito] ovvero con i <tag> HTML utilizzati dai webmaster per marcare le loro pagine in modo riconosciuto dai principali motori di ricerca come Bing, Google, Yahoo! e Yandex.

La mappatura tra alcuni elementi di "Schema.org" e metadati (dati bibliografici) che rappresentano le risorse di Biblioteche e Archivi digitali può migliorare significativamente l'esposizione e la ricerca delle medesime risorse sul Web.

Molti dati memorizzati in database relazionali e forniti da alcune applicazioni digitali sono spesso di poca utilità per i motori di ricerca Web, in quanto spesso questi motori falliscono ad interpretarli. Di conseguenza, gli utenti web hanno bassa probabilità di trovare i dati bibliografici sul web, a meno che essi non sappiano già esattamente dove le risorse risiedono, digitando i loro indirizzi web precisi.

Esplorando come i dati che rappresentano le risorse siano collegati tra loro sul Web, al fine di comprendere la struttura degli schemi in cui risiedono i medesimi dati, sarebbe opportuno prendere in considerazione non solo i grafi RDF e le strutture di *Linked Data* che collegano i dati [V. APPENDICE A], ma anche la marcatura dello "Schema.Org". Questo approccio è anche particolarmente rilevante per capire come migliorare le prestazioni delle applicazioni intelligenti per l'esposizione e la ricerca del materiale di ricerca sul web.

Diversi dati bibliografici sono stati già affiancati dai marcatori "Schema.Org" come, ad esempio, nei servizi di ricerca quali "WorldCat.org", "Data.bnf.fr", "VOA3R Open Access Repository" [V. siti].

Il motivo per cui "Schema.Org" è stato considerato nelle Raccomandazioni "LODE-BD" per la mappatura con gli elementi di metadati ha sostanzialmente due ragioni:

1) il vantaggio di creare micro-dati dalle singole fonti per esporli sul Web. Questo compito lo possono svolgere webmaster o autori stessi quando pubblicano i loro dati sul web, invece di passare attraverso un Repository per ottenere le esposizioni sul web. "Schema.Org" non sostituisce nessun schema di metadati, come, nel caso dello schema di metadati Dublin Core proposto dalle Raccomandazioni "LODE-BD" i cui elementi risultano complementari con i marcatori "Schema.Org";

2) lo "Schema.Org" è utilizzato da diversi agenti sul Web. Quindi molte proprietà dello "Schema.Org" impiegate per l'affinamento dei metadati nei Repository saranno utilizzate anche da altri agenti web. In questa prospettiva, l'interoperabilità tra le applicazioni sarà notevolmente potenziata. Repository possono anche raccogliere le risorse correlate dal web tramite l'impiego dei dati dallo "Schema.Org".

Il gruppo di lavoro "*Schema Bib Extend Community Group*" [V. sito] sta preparando diverse proposte per l'estensione dei vocabolari dello "Schema.Org" per migliorare la rappresentazione e ricerca dei dati bibliografici sul Web.

5.4. Uno Schema di metadati per la descrizione dei contenuti di ricerca negli Archivi Istituzionali

Nello schema di Metadati presentato nel presente paragrafo, in conformità alle indicazioni delle Linee guida CRUI, saranno fornite le mappature tra gli elementi di metadati Dublin Core (DC) con i gli elementi OAI_DC supportati dal protocollo dell'interoperabilità OAI-PMH (*Open Archive Initiative Protocol for Metadata Harvesting*).

Gli elementi di metadati proposte dalle Linee guida CRUI sono conformi alla codifica standard UNICODE, importante per la rappresentazione e la manipolazione di testo espresso nella

maggior parte dei sistemi di scrittura digitale. UNICODE utilizza lo schema XML come mezzo primario basato sul concetto "mix and match" per combinare gli elementi e sub-elementi, attributi correlati, e valori di attributi controllati in tutto il set di elementi.

Quanto al documento "LODE-BD", esso promuove la codifica di elementi di metadati all'interno di schemi RDF / XML per sostenere la loro coerenza semantica, necessaria nella maggior parte degli ambienti digitali. Sia le Linee guida CRUI che le Raccomandazioni "LODE-BD" assumono che i metadati da loro forniti possano essere più complessi e strutturati, prima di tutto per creare un quadro più equilibrato in grado di accomodare meglio diversi modelli di metadati in conformità alle diverse esigenze locali del Repository che concernono la rappresentazione e la gestione dei loro oggetti digitali.

L'obiettivo dell'allineamento di elementi dei metadati ricavati dalle Linee guida CRUI secondo le Raccomandazioni "LODE-BD" è quello di mostrare come elementi/termini di metadati, selezionati per la descrizione degli oggetti digitali negli Archivi Istituzionali possano essere migliorati/potenziati per mezzo della codifica offerta dagli Alberi decisionali "LODE-BD" nel presente Capitolo. L'allineamento dei metadati dei Repository Istituzionali con le metodologie di codifica offerte dalle Raccomandazioni "LODE-BD" porterà a:

1) una radicale trasformazione e ad un miglioramento dei flussi di lavoro (*workflows*) di metadati. L'integrazione delle loro proprietà con i valori "non letterali" di Vocabolari controllati e schemi di Classificazione farà risparmiare il tempo per lo sviluppo di nuovi e/o correlate indici di metadati;

2) una migliore descrizione e scoperta (tramite ricerca e navigazione) delle risorse degli Archivi entro il Web dei dati interconnessi (Web dei Dati, *Linked data*, Web 3.0).

I valori che abilitano i metadati a diventare *Linked Data* potranno reindirizzare gli utenti degli Archivi direttamente dalle interfacce di scoperta delle risorse ai diversi Sistemi di Gestione Dati DMS (*Data Management Systems*) che connettono diverse conoscenze sul Web Semantico;

3) l'integrazione dei dataset di Archivi entro la nuvola *Linked Open Data (LOD) Cloud* [V. sito]. In questo modo i contenuti dei *Repository* aumenteranno enormemente loro visibilità e integrazione sul Web Semantico;

4) un migliore scambio dei dati prodotti dai diversi soggetti produttori attraverso i dati collettivamente condivisi sulla base dei valori comuni *Linked Data*;

5) lo sviluppo dell'interfaccia di ricerca comune del tipo "*Institutional Repository WorldShare Platform*", come nel caso della piattaforma "*World-Cat local*" [V. sito] - un'esperienza di diversi "accessi multifocali" a 1.071 miliardi di risorse indicizzate dal tutto il mondo, a partire dal più grande Catalogo online delle collezioni delle biblioteche "*World-Cat*";

6) la creazione, la condivisione e l'uso di nuove applicazioni che possano potenziare i canali per la disseminazione e l'accessibilità delle entità informative espresse come *Linked Open Data* tramite i servizi di Archivi Istituzionali, contribuendo così alla creazione dello spazio virtuale per i contenuti di ricerca sul Web dei Dati - "*Open Research Commons Space*" [V. sito: *Building Institutional Repositories for Global Research Commons*; White, 2009, p.28].

Lo schema dei metadati che verrà presentato successivamente in Tabella 17 può essere considerato come un tentativo di creare un Profilo di applicazione (AP) di metadati per la descrizione delle risorse negli Archivi Istituzionali. Il concetto di AP è emerso nella comunità DCMI (*Dublin Core Metadata Initiative*) come un modo per dichiarare quali elementi da quali schemi di metadati sarebbe meglio utilizzare in una particolare Applicazione digitale o un Progetto di rappresentazione delle risorse. Questo significa che diversi elementi di metadati possono essere combinati insieme nei modi diversi, ottimizzando esigenze dei sistemi locali e descrittive dei servizi.

Lo schema presentato in Tabella 17 sarà basato sugli elementi di metadati proposti dalle Linee Guida CRUI e modellato secondo le Raccomandazioni "LODE-BD", riportando:

- 1) la strutturazione dei metadati in categorie secondo le Raccomandazioni "LODE-BD";
- 2) le strategie di codifica basate sull'uso motivato dei valori "non-letterali" (identificati con URIs) e "letterali" per la qualifica delle proprietà dei metadati;
- 3) le mappature tra i metadati Dublin Core e più specifici elementi di metadata, di Ontologia Bibliografica BIBO e dei marcatori dello "Schema.org".

Lo scopo dello schema di metadati presentato in seguito è anche quello di dimostrare l'utilità delle Raccomandazioni "LODE-BD" per migliorare la qualità espressiva dei metadati descrittivi *Dublin Core* entro gli Archivi Istituzionali.

Il medesimo schema può essere considerato come una parte della strategia "LODE-BD" "*Design-time*" per l'implementazione dei metadati. Entrambe le strategie - "*Design-time*" e "*Run-time*" saranno brevemente discusse dopo lo schema che segue nel prossimo paragrafo.

Gruppi di elementi / termini di metadati La scelta degli elementi di metadati deve essere fatta in base alle esigenze locali degli Archivi Istituzionali e alle strategie di codifica dei valori di metadati	Descrizione e commenti - Raccomandazioni per la compilazione - Strategie di codifica			
	<i>Cross-walks (mappature) a</i>			
	OAI_DC metadati	elementi di metadati più specifici	elementi dello "Schema.org" >: narrowMatch <:broadMatch	tipologie di marcatori dello "Schema.org" <i>punti indicano il livello di un sotto-tipo</i>
DC: DCTERMS:	Prefissi per namespace per specificare gli elementi di metadati Dublin Core DC: e DCTERMS:			
	<hr style="border-top: 1px dashed #000;"/> @prefix dc: <http://purl.org/dc/elements/1.1/> (DCMI <i>legacy namespace</i> usato per le proprietà di base dei 15 elementi DC)			
	<hr style="border-top: 1px dashed #000;"/> @prefix dcterms: <http://purl.org/dc/terms/> (DCMI <i>terms namespace</i> dei termini DCMI – proprietà, classi e tipi di dati)			
1. Title information	Descrizione e commenti: tramite questo Gruppo di metadati che comprende gli elementi "Proper Title"/ "Alternative Title"... - viene riconosciuta una Risorsa e attribuito un rilevante punto d'accesso alla medesima. Raccomandazioni per la compilazione: <ul style="list-style-type: none"> • Il nome del Titolo deve essere riportato come viene presentato nella Risorsa, iniziando con la lettera maiuscola. 			

- Se la risorsa definisce il Titolo in più lingue allo stesso tempo, ci dovrebbe essere un elemento autonomo con l'introduzione della/e marcatura/e della/e lingue.
- Se il titolo è assente, lo si può creare nel modo più descrittivo possibile, evitando i termini generici come, ad esempio, "Un articolo", "Una relazione" ecc.
- Il Sottotitolo deve essere salvato nell'elemento Titolo proprio ("*Proper Title*") utilizzando un formato di gap a trattino (, - ,) tra il Titolo e il Sottotitolo.
- Il punto alla fine del Titolo è omesso.
- L'elemento di *Alternative Title* deve essere usato solo nel caso in cui la risorsa è anche conosciuta con un nome diverso, compreso le abbreviazioni o acronimi.

Per il controllo di autorità della compilazione possono essere utilizzati i seguenti documenti:

1) RDA (*Resource Description and Access*) Standard:

<http://en.wikipedia.org/wiki/Resource_Description_and_Access>

2) Regole catalografiche nazionali specifiche (es. REICAT: le nuove regole di catalogazione italiana):

<http://www.iccu.sbn.it/upload/documenti/REICA_bozza_complessiva_genn2009.pdf>

Strategie di codifica:

1) proprietà dei termini di metadati - "*dc:title*", "*dcterms:title*" - "*dc:title.alternative*", "*dcterms.alternative*" - vengono qualificate solo con i valori "*letterali*" (stringhe di testo);

2) Sotto-titolo/i (*Subtitle(s)*) devono essere salvati nell'elemento "*dcterms:title*";

3) Titolo/i parallelo/i (*Parallel title(s)*) può/possono essere salvato/i sia in "*dcterms:title*" che in "*dcterms:alternative*";

4) Titolo/i tradotto/i (*Translated titles(s)*) devono essere salvati in "*dcterms:alternative*".

dc:title	dc:title		name	Thing
dcterms:title		dcterms:alternative	>headline	.CreativeWork
			>alternativeHeadline	.CreativeWork

<p>2. Responsible Body</p>	<p>Descrizione e commenti: Questa Categoria comprende tutti gli elementi di metadati che contengono le informazioni sugli agenti che esercitano la loro influenza sul contenuto della risorsa durante la creazione, la pubblicazione e la distribuzione della medesima, e può riferirsi a:</p> <ul style="list-style-type: none"> ✓ Creatore (<i>Creator</i>) = agente primario responsabile del contenuto intellettuale della risorsa. ✓ Contribuente (<i>Contributor</i>) = individuo (i) o organizzazione (i) e servizio (i) responsabili in modo secondario nel contribuire ai contenuti della risorsa. Nel caso in cui ci sono più Contribuenti, l'elemento può essere ripetuto riferendosi, ad esempio, a studenti, editori, revisori, coordinatori, tutor, lettori, critici, supervisori, redattori di atti di convegni, collezionisti di set di dati, traduttori e altri. ✓ Editore (<i>Publisher</i>) = l'entità (sia commerciale che non commerciale: Emittente, Produttore, Distributore, Proprietario) responsabile per la pubblicazione o la distribuzione della risorsa. Per i contenuti come Tesi, questo elemento deve contenere il nome dell'Università. <p>Raccomandazioni per la compilazione:</p> <ul style="list-style-type: none"> • L'elemento riferito all'organismo responsabile (<i>Responsible Body</i>) deve includere il nome dell'Autore, o dell'Ente. Questi Nomi possono essere definiti anche tramite gli identificatori URIs o tramite altri riferimenti che li descrivono. • Se il Software offre l'opportunità di inserire Cognome e Nome dell'organismo responsabile negli appositi spazi (box, caselle) separati, il nome dell'Ente deve essere compilato iniziando con la lettera maiuscola nello spazio apposito "Cognome". • Nel caso di una persona concreta - iniziare con la compilazione del suo Cognome con la lettera maiuscola iniziale. Dopo deve seguire il Nome proprio (completo o iniziali) o altri nomi. Es. "<i>Eco, Umberto</i>" o "<i>Eco, U.</i>" • Nel caso in cui ci sono più organismi responsabili (es. Creatori), l'ordine degli elementi dovrebbe riflettere la loro gerarchia formale, cioè il primo Autore, ad esempio, è considerato come il principale e gli altri - come i Coautori. Autori secondari devono essere salvati nell'elemento del metadato "<i>Contributor</i>". <p>Per il controllo di autorità della compilazione si raccomanda di utilizzati gli stessi documenti come per il controllo nel Gruppo di metadati "<i>Title information</i>".</p>
-----------------------------------	---

Creator	<p>Strategie di codifica:</p> <p>1) Verificare la disponibilità dei file di autorità (es. VIAF: <i>Virtual International Authority File</i>) che possono essere applicati per il controllo di autorità dell'identificazione dell'elemento del metadato "Creator".</p> <p>2) Verificare se i dati di autorità sono disponibili online come file <i>Linked Data</i>, denotando i suoi valori con gli identificatori URIs (valori "non letterali").</p> <p>3) In caso positivo, utilizzare i <i>namespace</i> per il metadato "dcterms:creator" per inserire gli appropriati valori "non letterali" (URIs).</p> <p>4) Le proprietà del metadato "dc:creator" potrebbero essere qualificate sia con il valore (in base alle esigenze locali dell'Archivio e strategie di codifica):</p> <ul style="list-style-type: none"> a) "letterale" che b) "non-letterale". 			
dc:creator	dc:creator	foaf:Person	creator	.CreativeWork
dcterms:creator			author	
Contributor	<p>Strategie di codifica:</p> <p>1) Verificare la disponibilità di File di autorità (es. VIAF: <i>Virtual International Authority File</i>) che possono essere applicati per il controllo di autorità dell'identificazione dell'elemento del metadato "Contributor".</p> <p>2) Verificare se i dati di autorità sono disponibili online tramite i file <i>Linked Data</i>, denotando i suoi valori con gli identificatori URIs (valori "non letterali").</p> <p>3) In caso positivo, utilizzare i <i>namespace</i> per il metadato "dcterms:contributor" per inserire gli appropriati valori "non letterali".</p> <p>4) Le proprietà del metadato "dc:contributor" potrebbero essere qualificate sia con il valore:</p> <ul style="list-style-type: none"> a) "letterale" e b) "non-letterale" <p>5) Valore/i "non letterale" (i) dell'organismo responsabile "Editor" può essere qualificato in modo incrociato con i valori dell'Ontologia BIBO (<i>namespace</i> "bibo:_")</p> <p>6) Proprietà riferite ai <i>namespace</i> del metadato "Contributor" possono essere qualificate anche in riferimento ai Ruoli forniti dallo schema Marcel (<i>MARC Code List for Relators</i>): <http://id.loc.gov/vocabulary/relators.html></p>			
dc:contributor.xx	dc:contributor		Contributor	.CreativeWork

dcterms:contributor			>editor	..Book
		bibo:editor	>Illustraor	
<i>Publisher</i>	Strategie di codifica: 1) Verificare la disponibilità di File di autorità (es. VIAF: <i>Virtual International Authority File</i>) che possono essere applicati per il controllo di autorità dell'identificazione dell'elemento del metadato " <i>Contributor</i> ". 2) Verificare se i dati di autorità sono disponibili online come i file <i>Linked Data</i> , denotando i suoi valori con gli identificatori URIs (valori "non letterali"). 3) In caso positivo, utilizzare i <i>namespace</i> per il metadato " <i>dcterms:publisher</i> " per inserire gli appropriati valori "non letterali". 4) Le proprietà del metadato " <i>dc:publisher</i> " potrebbero essere qualificate sia con il valore: a) "letterale" che b) "non-letterale".			
dc:publisher	dc:publisher	bibo:issuer	publisher	.CreativeWork
dcterms:publisher		bibo:producer	<provider	
		bibo:distributor		.CreativeWork
		bibo:owner		
3. Physical characteristics	Descrizione e commenti: Questa categoria deve essere descritta con riferimento a: <ul style="list-style-type: none"> ✓ Data (<i>Date</i>) = questo elemento del metadato compende informazioni temporali legate alla pubblicazione della risorsa. Quando questa voce non è disponibile, deve essere indicata una data in cui la risorsa è stata resa accessibile. ✓ Identificatore (<i>Identifier</i>) = serve per rendere efficace la localizzazione, il ritrovamento, lo scambio e il recupero di informazioni. Risorse e loro collezioni devono essere identificate dagli identificatori non ambigui univoci e persistenti all'interno di determinate strutture di riferimento (es. identificatori DOI, Handle che usano lo schema URI). ✓ Lingua (<i>Language</i>) = in questo elemento devono essere salvate le informazioni in riferimento alla/e lingua (e) utilizzata/e in connessione alla risorsa. ✓ Formato/Mezzo (<i>Format/Medium</i>) = in questo elemento deve essere specificato il formato di un oggetto digitale e il mezzo che rende disponibile il contenuto delle risorse. 			

	<ul style="list-style-type: none"> ✓ Edizione / Versione (<i>Edition/Version</i>) = edizione / versione / stato della risorsa. ✓ Fonte (<i>Source</i>) = è un riferimento alla fonte da cui proviene la risorsa primaria. Questo elemento può riferirsi a: pagina, volume, numero, riferimento, citazione nel testo e altre fonti.
<i>Date</i>	<p>Raccomandazioni per la compilazione:</p> <p>La compilazione dell'elemento del metadato Data (<i>Date</i>) deve essere in linea con lo Standard W3CDTF (World Wide Web Consortium Date Time Format) ISO 8601, 1997: <http://www.w3.org/TR/NOTE-datetime>. Gli esempi di formati della compilazione:</p> <ul style="list-style-type: none"> • YYYY; YYYY-MM; YYYY-MM-DD; YYYY-MM-DDThh:mmTZD (es. 2013-12-30T08:30+1:00); • YYYY-MM-DDThh:mm:ssTZD (es. 2013-12-30T08:30:25+1:00); • un formato completo: YYYY-MM-DDThh:mm:ss.sTZD (es. 2013-09-30T08:30:25.45+1:00); <p>Le risorse per le quali può essere rilevante il metadato riferito alla data di accettazione "<i>dc: dateAccepted</i>" sono:</p> <ul style="list-style-type: none"> • una Tesi (accettata dall'unità protocollo dell'Università) o • un Articolo (accettato da una Rivista). <p>Nel caso in cui la data formale di pubblicazione di una risorsa è complessa (es. YYYY-MM-DD), nella compilazione dell'elemento "<i>dc:date.Issued</i>" - solo l'elemento Anno (YYYY) viene considerato un elemento indispensabile / obbligatorio.</p> <p>Strategie di codifica</p> <p>1) È una buona pratica selezionare e seguire una sintassi di codifica dettata da uno Standard o proposta da determinate Linee guida.</p> <p>2) Per codificare le tipologie di date non differenziate (senza ulteriori estensioni di specificazione), possono essere utilizzati i seguenti <i>namespace</i> (ambidue o in alternativa):</p> <ol style="list-style-type: none"> a) "<i>dc:date</i>" (solo il "valore letterale"; es. 2013); b) "<i>dcterms:date</i>" (solo il "valore letterale"; es. 2013-04; 2013-04-14)

	3) Se le date vengono ulteriormente specificate (es. <i>Issued, Accepted</i>), è consigliato utilizzare i <i>namespace</i> di metadati " <i>dcterms: xx</i> " (es. " <i>dcterms:date.issued</i> ") codificati con i "valori letterali" (come secondo le Raccomandazioni "LODE-BD").			
dc:date	dc:date	dcterms:created	dataCreated	.CreativeWork
dcterms:date		dcterms:dateCopyright	copyrightYear	
		dcterms:modified	dataModified	
		dcterms:available	dataPublished	
		dcterms:dateSubmitted		
		dcterms:dateAccepted		
		dcterms:valid		
		dc:description.conferenceDate		
		Dc:date.embargoEnd		
Note:				
<ul style="list-style-type: none"> ✓ Il <i>namespace</i> per l'elemento di metadati "<i>dc:description.conferenceDate</i>" potrebbe essere importante in riferimento ad alcune risorse degli Archivi Istituzionali. ✓ L'elemento "<i>dc:date.embargoEnd</i>" per il singolo oggetto digitale deve essere appropriatamente trasmesso secondo lo Standard PEG21 DIDL che definisce sia un modello astratto che uno schema XML per rappresentare entità di oggetti complessi. ✓ La presenza dell'elemento "<i>dc: date.embargoEnd</i>" è altamente raccomandato per la descrizione dei prodotti di ricerca realizzati nell'ambito del 7 ° (FP7) Programma Quadro, depositati nei Repository compatibili con i requisiti (Linee Guida) definiti dal Progetto OpenAIRE: https://www.openaire.eu/ ✓ La tipologia dei dati per l'elemento <code><dc:date.embargoEnd></code> è controllata da <i>namespace</i> SURF: <code><info:eu-repo/date/embargoEnd/></code> (<code><http://wiki.surf.nl/display/standards/info-eu-repo/></code>). Codifica di questo valore deve essere eseguita in conformità con la forma YYYY-MM-DD (Standard ISO 8601). ✓ Il valore per l'elemento "<i>dc: dateSubmitted</i>" deve essere inserito automaticamente dal sistema al termine della procedura di deposito (<i>datastamp</i>). 				
Identifier	Raccomandazioni per la compilazione:			

• Nelle descrizioni bibliografiche, una risorsa deve essere sempre rappresentata da un identificatore ID univoco. Questo ID può essere assegnato a livello locale (o essere locale temporaneamente), o essere lo stesso come i suoi identificatori globali quali: *Uniform Resource Identifier / URI*, *Uniform Resource Locator / URL*, *Uniform Resource Name / URN*, o contiene la stringa da un identificatore universale come *Digital Object Identifier / DOI*, *International Standard Book Number / ISBN*, *Handle*.
V. sito: <<http://www.openarchives.org/OAI/openarchivesprotocol.html#UniqueIdentifier>>

Strategie di codifica:

Nel trattamento degli identificatori (globali o locali) si consiglia sempre di verificare la loro sintassi seguendo specifici Standard e/o Linea Guida.

- 1) Se non ci sono tipi di identificatori differenziati, salvarli (in alternativa o insieme, in base alle scelte fatte per l'Archivio) in *namespace*: "*dc: identifier*", "*dcterms: identifier*" (utilizzabili solo con i valori "*letterali*").
- 2) Se ci sono tipi di identificatori differenziati, salvarli anche in *namespace* dell'Ontologia BIBO ("*Bibo: _*")
- 3) Per specificare gli elementi di metadati che descrivono Articoli, Blog, Libri e altri contenuti editoriali online, possono essere usati i *namespace* dei metadati PRISM: "*prism: _*" (<<http://prismstandard.org/namespaces/basic/2.1/>>) (es. "*prism:doi*"; "*prism:issn*"; "*prism:eissn*").

dc:identifier	dc:identifier	bibo:asin		
dcterms:identifier		bibo:coden		
		bibo:eanucc13		
		bibo:eissn		
		bibo:gtin14		
		bibo:handle		
		bibo:isbn	isbn	..Book
		bibo:issn		
		bibo:lccn		
		bibo:oclcnum		
bibo:pmid				

		bibo:sici		
		bibo:upc		
		bibo:uri		
		bibo:locator		
Language	<p>Raccomandazioni per la compilazione:</p> <ul style="list-style-type: none"> • Se ci sono più lingue utilizzate in riferimento alla risorsa, l'elemento <i>Language</i> deve essere ripetuto per ogni lingua. • Per l'identificazione della lingua è consigliabile utilizzare una lista di codici controllata. Per esprimere il codice di lingua in due lettere (es. "en" per l'inglese) – utilizzare lo Standard ISO 639-1. Per esprimere il codice di lingua in tre lettere (es. "eng" per l'inglese) – utilizzare lo Standard ISO 639-2. <p>Nel caso di una traduzione, deve essere indicate anche la lingua del lavoro originale.</p> <p>Strategie di codifica:</p> <p>Salvare identificatori della lingua (in modo alternativo o assieme) nei seguenti <i>namespace</i> di metadati:</p> <ol style="list-style-type: none"> a) <i>dc:language</i> (con valore "letterale"); b) <i>dc:language</i> (con valore "non-letterale"); c) <i>dcterms:language</i> (con valore "non-letterale" URI che può abilitare il meccanismo di <i>Linked Data</i>). 			
dc:language	dc:language		inLanguage	.CreativeWork
dcterms:language				
Format/Medium	<p>Raccomandazioni per la compilazione:</p> <ul style="list-style-type: none"> • Si consiglia sempre di utilizzare un vocabolario controllato, come ad esempio la lista di <i>Internet Media Types</i> [MIME: <http://www.iana.org/assignments/media-types>]; formati PDF, text/HTML, text/XML. • Si raccomanda di utilizzare un Vocabolario controllato per la descrizione della collezione quando si usa <i>namespace</i> <i>dcterms:medium</i> per definire il materiale o il supporto fisico della risorsa. <i>Internet Media Types</i> [MIME] non 			

	<p>deve essere utilizzato per supportare questi valori.</p> <ul style="list-style-type: none"> • In caso di presenza di più formati legati a una risorsa, questi possono avere differenti politiche di accesso. <p>Strategie di codifica:</p> <p>1) Se non vi è alcuna lista di codici controllata o se non c'è un vocabolario abilitante <i>Linked Data</i> disponibile (con i valori "non-letterali" URIs), codificare i valori nello spazio di namespace "<i>dc:format</i>" (con il valore "letterale").</p> <p>2) Se esiste un elenco di codici, e se esiste un vocabolario abilitante <i>Linked Data</i> disponibile (con i valori "non-letterali" URIs), codificare i valori in namespace: "<i>dcterms: format</i>", "<i>dcterms: medium</i>" (con i valori "non letterali"), in alternativa o in ambedue <i>namespace</i>.</p>			
dc:format	dc:format	dcterms:medium		
dcterms:format				
<i>Edition/Version/Status</i>	<p>Raccomandazioni per la compilazione:</p> <ul style="list-style-type: none"> • Controllare se la versione dell'edizione è definita da alcune informazioni. • Informazioni sul "<i>Edition</i>" è obbligatoria per i libri in edizione successiva (es. 2.ed., 2.ed.riv.e corr., V.3.0). <p>Strategie di codifica:</p> <ul style="list-style-type: none"> • Se non ci sono tipologie specifiche differenziate, codificare valore (i) in <i>namespace</i>: "<i>dc:description</i>", "<i>dcterms:description</i>" (con i valori "letterali"), in alternativa o in ambedue <i>namespace</i>. • Se vi sono tipologie specifiche differenziate, codificare valore (i) in <i>namespace</i> dell'Ontologia BIBO: "<i>bibo:edition</i>", "<i>bibo:status</i>" • Se devono essere descritte relazione (i) tra una risorsa e sua/e versione/i correlata (e), vedere la sezione del Gruppo dei metadati "<i>Relation</i>" ("<i>dc:relation</i>", "<i>dcterms:hasVersion</i>", "<i>dcterms:isVersionOf</i>", "<i>dcterms:replaces</i>", "<i>dcterms:isReplacedBy</i>"). • Secondo le "Linee guida CRUI" può essere utilizzato anche il <i>namespace</i> "<i>prism:edition</i>". 			
dc:description.edition	dc:description	bibo:edition	>bookEdition	..Book
		bibo:status	>version	..CreativeWork

Source	<p>Raccomandazioni per la compilazione:</p> <ul style="list-style-type: none"> • I valori di questo elemento devono contenere i dati riferiti alla locazione e altre descrizioni della fonte “<i>Source</i>”. • L’elemento “<i>Source</i>” può anche offrire uno spazio per la codifica della citazione bibliografica; • Se il nome della fonte come “<i>Title</i>” è controllato tramite un <i>Authority File</i>, utilizzare il Titolo controllato o un Identificatore. • È raccomandato l’utilizzo degli schemi BibTex, APA, MLA. • Le Linee Guida <i>OpenAIRE Guidelines 2.0</i>. raccomandano di riferirsi a <i>OpenURL ContextObject</i> dalle Linee Guida per l’informazione citazionale bibliografica per la codifica definite da <i>Dublin Core Metadata DCMI</i>: <http://dublincore.org/documents/dc-citation-guidelines/> <p>Strategie di codifica:</p> <p>1) Controllare se il riferimento alla risorsa è contenuta in un’altra risorsa. Codificare il/i riferimento (i) in <i>namespace</i>:</p> <p>a) “<i>dc: source</i>” (con i valori "letterali") (es. “<i>Proceedings of the National Academy of Sciences of the United States of America, v. 95(10) p. 5632-5636</i>”);</p> <p>b) “<i>dcterms:source</i>” ("non letterale").</p> <p>2) Se ci sono i titoli separati da diverse fonti - codificare i loro valori in <i>namespace</i>: “<i>dc:source</i>” (“ letterale”), “<i>bibo:_</i>” ("letterale"), in alternativa o in entrambi <i>namespace</i>.</p> <p>3) In accordo con le "Linee guida CRUI", i valori possono essere anche codificati in <i>namespace</i>: “<i>prism:volume</i>”, “<i>prism:number</i>”, “<i>prism:startingPage</i>”, “<i>prism:endingPage</i>”.</p> <p>4) Per la/e relazione (i) tra una risorsa e le sue fonti correlate, vedere la sezione del Gruppo dei metadati “<i>Relation</i>” (“<i>dc:relation</i>”; “<i>dcterms:hasPart</i>”, “<i>dcterms:isPartOf</i>”, “<i>bibo:presentedAt</i>”, “<i>bibo:reproducedIn</i>”, “<i>bibo:documentPart</i>”).</p>			
dc:source	dc:source	bibo:pages	>numberOfPage	..Book
dcterms:source		bibo:pageStart		
dc:identifier	dc:identifier	bibo:pageEnd		
dcterms:identifier		bibo:section		

		bibo:volume	bibo:issue	
		bibo:chapter		
4. Holding/ Location Information	<p>Descrizione e commenti:</p> <p>Per la sicurezza delle informazioni sulla locazione e sulla disponibilità di identificazione delle risorse per lo scambio di informazioni, l'uso di questo elemento di metadati è essenziale (anzi obbligatorio).</p> <p>Raccomandazioni per la compilazione:</p> <p>È sempre consigliabile che le informazioni sulla localizzazione siano fornite in un modo consistente seguendo una regola di codifica o di Linee Guida.</p> <p>Strategie di codifica:</p> <p>1) Controllare se la risorsa ha le informazioni sulla propria localizzazione. Altrimenti, individuare o assegnare una localizzazione alla risorsa.</p> <p>2) Controllare se le informazioni sulla localizzazione della risorsa seguono una specifica regola di codifica o Linee guida.</p> <p>3) Codificare valori sulla localizzazione in <i>namespace</i> (in alternativa o in entrambi <i>namespace</i>):</p> <p>a) “<i>agls:availability</i>” (valori “letterali”, es. “<i>University of Vienna, Peter Jordanstr. 52, A-1190 Vienna, Austria</i>”),</p> <p>b) “<i>bibo:locator</i>” (valori “letterali”, es. “Box 12, Folder 3”).</p>			
		agls:availability	ContentLocation	.CreativeWork
		bibo:locator		
5. Subject	<p>Descrizione e commenti: Questo elemento serve a descrivere l'argomento della risorsa (di ciò che la risorsa denota) tramite parole chiave e/o frasi chiave.</p> <p>Raccomandazioni per la compilazione:</p> <ul style="list-style-type: none"> • Secondo le Linee guida <i>OpenAIRE V 2.0</i>, le parole chiave e/o i codici/i termini dalle Classificazioni possono essere di testo libero ovvero "letterali" (preferibilmente in lingua inglese) e definiti dagli identificatori web URIs (per 			

	<p>l'abilitazione di <i>Linked Data</i>). (Le Linee consigliano di fare il riferimento al <i>namespace</i> <info:eu-repo/classification>).</p> <ul style="list-style-type: none"> • Per specificare meglio alcuni contenuti di ricerca - come ad esempio "Tesi" – possono essere applicati gli schemi specifici quali: <ul style="list-style-type: none"> - Settori Scientifico Disciplinari (SSD), - CRUI schema (per il contesto italiano) [PARTE II, Cap. 4, §§ 4.3., 4.4] codificando i loro valori "letterali" in <i>namespace</i> “<i>dc:subject.ssd</i>”, “<i>dc:subject.isicrui</i>”. • V. la Tabella 15. “Alcuni Vocabolari controllati e Schemi semantici per la normalizzazione e l'abilitazione dei metadati d'Archivio come <i>Linked Data</i>“ • V. l'elenco di <i>Value Vocabularies</i> disponibili online come <i>Linked Open Data</i>, con i relativi casi d'uso: <http://www.w3.org/2005/Incubator/lld/XGR-lld-vocabdataset-20111025/> • L'utilizzo di Sistemi di Organizzazione della Conoscenza (KOS) con i valori "non-letterali" (URI) abilitanti i meccanismi <i>Linked Data</i> offrono un reale potenziamento della qualificazione dei valori delle proprietà del metadato "Soggetto". <p>Strategie di codifica:</p> <ol style="list-style-type: none"> 1) Controllare se la risorsa ha uno o più Soggetto/i o argomento/i . 2) Controllare se è disponibile un vocabolario controllato per la descrizione dei valori del “Soggetto”: <ol style="list-style-type: none"> a) Se “No” – salvare l'informazione disponibile sul Soggetto in <i>namespace</i> dell'elemento di metadato “<i>dc:subject</i>” (valore “letterale”) b) Se “Si” - controllare se il vocabolario disponibile definisce i propri valori tramite valori “non-letterali” (URIs), abilitando così il meccanismo di pubblicazione <i>Linked Data</i>. c) Se ci sono le tipologie specifiche differenti per descrivere un Soggetto, codificare le relative informazioni in <i>namespace</i> (in alternativa o in ambedue <i>namespace</i>): <ul style="list-style-type: none"> - “<i>dc:subject</i>”, - “<i>dc:coverage</i>” (valore “letterale”);
--	---

	<p>d) Se è stato trovato un vocabolario con i valori “non-letterali” (URIs) per l’abilitazione di <i>Linked Data</i>, ma non ci sono diversi tipologie di voci per descrivere un argomento - salvare le informazioni in namespace: “<i>dcterms:subject</i>” (valori "non letterali", URIs);</p> <p>e) se sono presenti sia il vocabolario abilitante <i>Linked Data</i> che diverse tipologie delle voci descriventi uno specifico argomento, salvare le informazioni in <i>namespace</i> (in alternativa o in tutti): “<i>dcterms:subject</i>”, “<i>dcterms:coverage</i>”, <i>dcterms:spatial</i>”, “<i>dcterms:temporal</i>” (valori "non letterali", URIs).</p>			
dc:subject	dc:subject		About	.CreativeWork
dcterms:subject			Keywords	
dc:coverage	dc:coverage		Mentions	
dcterms:coverage			dcterms:spatial	
		dcterms:temporal	Keywords	
6. Description of content	<p>Descrizione e commenti: Questo elemento caratterizza la descrizione di due categorie principali: 1) descrizione del contenuto (es. Descrizione, Abstract, Bibliografia, Indice, Sintesi, Note); 2) descrizione del tipo (forma) o del genere della risorsa.</p> <p>Raccomandazioni per la compilazione:</p> <ul style="list-style-type: none"> • Compilare l’Abstarct (l’interpretazione di <i>default</i> per il valore del metadato "<i>dc: description</i>"), utilizzando lo Standard “ISO 214:1976: <i>Documentation - Abstracts for publications and documentation</i>”. • Abstarct in Inglese <p>Strategie di codifica:</p> <ol style="list-style-type: none"> 1) Controllare se la risorsa ha una descrizione (Abstract/ Note / Sommario / Descrizione). 2) Controllare se la descrizione ha diversi tipi specifici per la sua definizione: <ol style="list-style-type: none"> a) Se "No" - salvare le informazioni in <i>namespace</i> dell’elemento del metadato “<i>dc:description</i>” b) Se "Si" - codificare i valori in <i>namespace</i> (in alternativa o in tutti): “<i>dcterms:abstract</i>”, “<i>dcterms:tableOfContents</i>”, “<i>dcterms:description</i>” (che accettano entrambi i valori: "non-letterali" e "letterali"). 			
dc:description	dc:description			

dcterms:description		dcterms:tableOfContent	description	
dc:abstract	dc:abstract			
dcterms:abstract				
<i>Type/Status/Genre</i>	<p>Descrizione e commenti: Questo elemento caratterizza la descrizione della relazione tra una risorsa e il suo tipo/stato(forma)/ genere.</p> <p>Raccomandazioni per la compilazione:</p> <ul style="list-style-type: none"> • Si consiglia sempre di utilizzare un vocabolario controllato. • Linee guida CRUI [V. sito, pp.13, 14) offrono un elenco di tipologie dei documenti che possono essere depositati negli Archivi Istituzionali italiani, riportando anche la loro mappatura con le tipologie documentali definiti da SURF. Questo elenco comprende: articolo, annotazione, revisione, Tesi di Dottorato, Tesi di Laurea (triennale, magistrale), <i>Working Paper</i>, libro, parte del libro, Oggetto di Conferenza, Articolo di Conferenza, Poster di Conferenza. • Nel 2012 MIUR ha definito l'elenco dei tipi di pubblicazione selezionabili per la VQR (Valutazione della Qualità della Ricerca): <http://hubmiur.pubblica.istruzione.it/web/ricerca/valutazione> • Stato/Tipologia della Versione di una pubblicazione (es. "articolo") è uno stato raggiunto dalla risorsa nel suo processo di flusso di lavoro di pubblicazione (nel suo flusso generale e non relativo al deposito nell'Archivio Istituzionale). • Per qualificare gli elementi di stato, fare riferimento a <i>namespace</i> della fondazione SURF : <http://wiki.surffoundation.nl/display/standards/info-eu-repo#info-eu-repo-Versions> • Linee guida OpenAIRE V 2.0 raccomandano di definire tipo di pubblicazione e il tipo della Versione tramite uno sistema di notazione URI (preferibilmente utilizzare <i>namespace</i> : <info:eu-repo/semantics>). <p>Strategie di codifica:</p> <ol style="list-style-type: none"> 1) Controllare se la risorsa ha Tipo / Forma / Genere / Stato. 2) Controllare se è disponibile una o piu' liste controllate di codici <ol style="list-style-type: none"> a) Se "No" – codificare i valori in <i>namespace</i> di elementi dei metadati “<i>dc:type</i>”, “<i>dc:format</i>” (valori 			

	<p>"letterali").</p> <p>b) Se "Si" – controllare se disponibile un vocabolario controllato che abilita i meccanismi di <i>Linked Data</i> tramite i valori "non-letterali" (URIs),</p> <p>b.1) Se "No" - codificare i valori in <i>namespace</i> (in alternativa o in ambedue): "<i>dc:type</i>", "<i>dc:format</i>" (valori "letterali").</p> <p>b.2) Se "Si" - codificare i valori in <i>namespace</i> (in alternativa o in tutti): "<i>dcterms:type</i>" "<i>dcterms:format</i>", "<i>dcterms:medium</i>" (con valori "non-letterali").</p>			
dc:type	dc:type		Genre	.CreativeWork
dcterms:type		dcterms:type.xxx		
dc:format	dc:format			
dcterms:format		dcterms:medium		
7. Intellectual Property Rights	<p>Descrizione e commenti: Questo elemento è utilizzato per trasmettere le informazioni sui diritti della risorsa e sul suo proprietario, che dovrebbe essere il detentore dei diritti sulla risorsa. Le proprietà dei metadati collegate con la questione dei diritti di proprietà sull'opera sono: <i>Rights</i> o <i>Rights Statement</i>. Le dichiarazioni più dettagliate possono essere codificate nelle seguenti proprietà di metadati: <i>Access Rights</i>, <i>Terms Of Use</i>, <i>Access condition/Access Rights</i>, <i>License</i>.</p> <p>Raccomandazioni per la compilazione:</p> <ul style="list-style-type: none"> • Si raccomanda l'uso dello Standard <i>Date Time Format</i> ISO 8601. • Per descrivere gli oggetti di ricerca in modalità di accesso <i>Open Access</i> è ampiamente accettato l'utilizzo della Licenza <i>Creative Commons Open License</i>. • V. Tabella 16: Alcuni vocabolari controllati per l'abilitazione dei metadati <i>RightsHolder</i>, <i>Licence</i>, <i>AccessRights</i>, <i>Rights</i> come <i>Linked Data</i>. • La compilazione di metadata sulle dichiarazione dei diritti per i prodotti di ricerca realizzati nell'ambito del 7 ° (FP7) Programma Quadro, depositati entro i Repository compatibili con i requisiti (Linee Guida) definiti dal Progetto OpenAIRE: <https://www.openaire.eu/> e' obbligatorio l'uso di seguenti <i>namespace</i>: "<i>dc.rights.accessmode</i>", "<i>dc.rights.embargo_end_date</i>". <p>Strategie di codifica:</p> <p>1) Controllare se la risorsa ha dichiarazioni sui diritti / condizioni / termini di utilizzo e di accesso.</p>			

	<p>2) Controllare se le medesime dichiarazioni sono dei diversi tipi.</p> <p>a) Se “No” – codificare i valori in <i>namespace</i>: “<i>dc:type</i>”, “<i>dc:format</i>” (con valori “letterali”).</p> <p>b) Se “Si” - controllare se disponibile un vocabolario controllato che abilita i meccanismi di <i>Linked Data</i> tramite i valori “non-letterali” (URIs):</p> <p>b.1) Se “No” - codificare i valori in <i>namespace</i> (in alternativa o in ambedue): “<i>dc:rights (letters)</i>”, “<i>dcterms:rights</i>” (“non-letterali”);</p> <p>b.2) Se “Si” - codificare i valori in <i>namespace</i> (in alternativa o in tutti): “<i>dcterms:accessRights</i>”, “<i>dcterms:licence</i>”, “<i>dcterms:rightHolder</i>”, “<i>dcterms:rights</i>” (“non-letterale”).</p>		
<p>dc:rights</p> <p>dcterms:rights</p>	<p>dc:rights</p>	<p>dc:rights</p> <p>dcterms:rights</p> <p>dc:accessRights</p> <p>dc:rights.accessMode</p> <p>dc:rights.embargo_end_date</p>	<p>copyrightHolder</p> <p>.CreativeWork</p>
<p>8. Relations</p>	<p>Descrizione e commenti: le relazioni possono descrivere le risorse da diverse prospettive. Elemento di relazione può riguardare o fare riferimento a:</p> <ul style="list-style-type: none"> - risorsa intesa come una Rivista o una Collezione, Sotto-Collezione (es.: “<i>dcterms: hasPart</i>”, “<i>dcterms: isPartOf</i>”), - altre risorse per mezzo di riferimenti rilevanti (es. una stringa o un numero conforme ad una identificazione formale della risorsa: identificatori DOI, Handle URL entro lo schema URI). <p>Le relazioni possono essere anche utilizzate per identificare le connessioni tra diversi agenti, versioni, traduzioni (<i>agents, versions, translations</i>) ecc.</p> <p>I dati salvati nella categoria dei metadati “<i>Relations</i>” migliorano le prestazioni del sistema di ricerca dei contenuti. Una Relazione qualificata con i valori “non-letterali” è un elemento molto rilevante per trasformare un Archivio Istituzionale “statico” in un Archivio dinamico per l’abilitazione dei <i>Linked Data</i>.</p> <p>Raccomandazioni per la compilazione:</p> <ul style="list-style-type: none"> • Controllare se è disponibile un File di Autorità per identificare i nomi degli agenti (es. VIAF: <i>Virtual International Author File</i>). 		

• E' altamente raccomandato che gli agenti siano sempre rappresentati dai loro identificatori univoci con i valori controllati:

• Le relazioni tra agenti possono essere espresse tramite:

- lo schema MarcRel: <marcrel:FND> dove, ad esempio, "Funder" ha un URI: <<http://id.loc.gov/vocabulary/relators/fnd.html>> e può essere rappresentato dal nome controllato "The Mellon Foundation";
- il *Getty Thesaurus of Geographic Names* ® Online, i valori di cui possono qualificare le proprietà dell'elemento "Place" al quale può essere connessa una risorsa;
- il profilo applicativo (AP) per le risorse di ricerca SWAP (*Scholarly Works Application Profile*) (es. il valore "University of Bristol" salvato in *namespace* dell'elemento del metadato "eprint:affiliatedInstitution");
- elementi "dc:relation.project", "eprint:grantNumber" che devono contenere un numero di un *Grant Agreement*. La presenza di questi elementi entro il *namespace* <info:eu-repo/grantAgreement/EC/FP7/PITN-GA-2009-237252>) è obbligatorio nella descrizione dei prodotti di ricerca nell'ambito del 7 ° (FP7) Programma Quadro, depositati nei Repository compatibili con i requisiti (Linee Guida) definiti dal Progetto OpenAIRE: <<https://www.openaire.eu/>>

Il numero di un *Grant Agreement* può essere salvato in *namespace* del profilo applicativo SWAP (es. A456X nell'elemento "eprint:grantNumber").

• Le Linee Guida OpenAIRE Guideines (V.2.0) hanno introdotto un *namespace* esteso per esprimere in modo generico l'identificazione del progetto (*ProjectID*) che può equivalere all'identificativo del *Grant Agreement*.

La compilazione dei valori in *namespace*

<info:eu-repo/grantAgreement/Funder/FundingProgram/ProjectNumber>

è obbligatorio per collegare le informazioni sul Progetto alla pubblicazione nello spazio informativo delle piattaforme compatibili con i requisiti definiti dall'OpenAIRE per le piattaforme di ricerca. L'uso di altri *namespace* obbligatori è definito nelle Linee Guida OpenAIRE Guideines (V.1.1).

Strategie di codifica:

- 1) Controllare se la Risorsa è correlata ad altra/e risorsa/e.
- 2) Se "Si", prendere una decisione in merito alla strutturazione e alla descrizione delle eventuali relazioni.
- 3) Controllare se ci sono diverse tipologie delle relazioni tra le risorse.
 - a) Se "No" – codificare il valore identificativo della risorsa in *namespace* (in modo alternativo o in entrambi):
 - a.1) "dc:relation" (accetta entrambi i valori "letterali" e "non-letterali")

	<p>a.2) “<i>dcterms:relation</i>” (“non-letterali”).</p> <p>b) Se “Si” - codificare i valori in <i>namespace</i> (in modo alternativo o in tutti): (in accordo con la scelta fatta per i metadati e le strategie di codifica per i valori “letterali” e “non-letterali”), prendendo in considerazione l’Albero decisionale delle Raccomandazioni “LODE-BD” - 3.9. <i>Decision Tree</i> “<i>Relation</i>”.</p>		
dc:relation	dc:relation	dc:relation.conference	
dcterms:relation		dc:relation.conferencePlace	
		dc:relation.event	
		dc:relation.eventdate	
		dc:relation.book	
		dc:relation.edition	
		dc:relation.series	
		dc:relation.journal	
		dc:relation.journal.code	
		dc:relation.volume	
		dc:relation.issue	
		dc:relation.place	
		dc:relation.project	
		dc:isReferencedBy	
		dc:references	
		dcterms:isRequiredBy	
		dcterms:requires	
		bibo:annotates	
		bibo:citedBy	
		bibo:cites	
		dcterms:identifier	
		bibliographiccitation	
		dcterms:hasFormat	
		dcterms:hasVersion	
		dcterms:isVersionOf	

		dcterms:hasPart		
		dcterms:isPartOf		
		dcterms:isReplacedBy		
		dcterms:replaces		
		dcterms:hasTranslationOf		
		dcterms:isTranslationOf		
		bibo:translationOf		
		dcterms:hasMetadata		
		dcterms:isShownBy		
		dcterms:isShownAt		
		eprint.affiliatedInstitution	affiliation	.Person
			>sourceOrganization	.CreativeWork
		marcel:FND	>sponsor	.MedicalEntity
				.MedicalStudy
		eprint:grantNumber		
		dc:relation.projectID		
L'uso dei metadati “ <i>isShownBy</i> ” – “ <i>isShownAt</i> ”, “ <i>citedBy</i> ” - “ <i>cites</i> ”, “ <i>isRequiredBy</i> ”- “ <i>requires</i> ”, “ <i>isReplacedBy</i> ”- “ <i>replaces</i> ”... è importante per distinguere in modo esplicito tra i riferimenti diretti e indiretti alla risorsa.				

Tabella 17. Schema di metadati per la descrizione dei contenuti di ricerca negli Archivi Istituzionali basato sulle Linee Guida CRUI e le Raccomandazioni “LODE-BD”

5.4.1. Strategie di implementazione dei metadati sviluppati alla luce di "LODE-BD"

Per implementare i metadati elaborati secondo le strategie "LODE-BD", i fornitori di dati (*Data Provider*, Archivi di dati) possono seguire due successive opzioni [V. sito: *Linked Open Data-enabled Bibliographical Data Recommendations*, p.44] (Figura 57):

1. "*Design-time*". Questa strategia prevede la sostituzione del modello di metadati corrente, sviluppando un altro schema secondo le proposte "LODE-BD" per la selezione e la modellazione dei metadati descrittivi. La scelta di questa strategia significa anche apportare alcune modifiche al Database di metadati correnti e ai Servizi che lo accedono;
2. "*Run-time*" (o "*on the fly*") [V. sito: *National Science Digital Library* (NSDL)].
Questa opzione significa che - pur mantenendo invariati il modello dei metadati corrente e la struttura del database - i *Data Provider* dovrebbero aggiungere i servizi di conversione per abilitare una mappatura tra i metadati dell'Archivio e di tradurre così i valori di metadati da "letterali" a "non letterali" (identificati con gli indirizzi URIs), seguendo le strategie di codifica degli "Alberi Decisionali" LODE-BD.

In particolar modo, il modello di integrazione "on the fly" potrebbe essere realizzato attraverso l'integrazione dell'Archivio (*Repository*) con un *Content Management System* (CMS) – DRUPAL [V. sito] - un *Open Source* dall'architettura modulare molto diffuso che permette di costruire vere e proprie applicazioni Web integrate con i servizi esterni. I dati "on the fly" saranno letti "al volo" dall'Archivio e reindirizzati per la loro visualizzazione in formato XHTML per il browser.

DRUPAL non offre intanto uno strato di archiviazione delle risorse digitali; questo significa che le risorse non diventano nodi di DRUPAL. In secondo luogo, la lettura dei dati "on the fly" non sarà diretta dal *Repository*. Ad esempio, essa potrebbe avvenire dialogando con lo strato intermedio tra il *Repository* e la piattaforma DRUPAL - Service Delivery Platform (SDP) [V. sito] che espone una serie di servizi REST [V. sito]. Tramite questi servizi, la piattaforma del *Repository* può "venire a conoscenza" della struttura dei dati/contenuti che "vogliono" interagire con il *Repository* e così può essere effettuato *rendering*¹ che compone il *layout* del sito DRUPAL con i metadati (combinati nelle schede descrittive di dettaglio: metadati sommari ed estesi) delle risorse digitali del *Repository*.

Nella Figura 57 dovuta attenzione viene prestata alla rappresentazione della strategia "Run-Time", puntando sulla conversione del valore del metadato "Soggetto" dal "letterale" al "non letterale" (identificato con URI). Come è dimostrato, l'informazione sul "Soggetto" può essere descritta sia dal valore "letterale" ("Giappone", dal *namespace* di *Dewey Classification language @dew*) che dal valore "non letterale" (URI che esprime il concetto equivalente dal Servizio *Linked Data* "Dewey.info").

Entrambi i valori "letterali" e "non letterali" possono essere tracciati come i grafi RDF (triple "Soggetto/Risorsa – Proprietà/Relazione – Oggetto/Valore" incrociate) salvati in *TripleStore* (Figura 58) integrati nelle piattaforme degli Archivi implementati sui Software semantici (che supportano *TripleStore*).

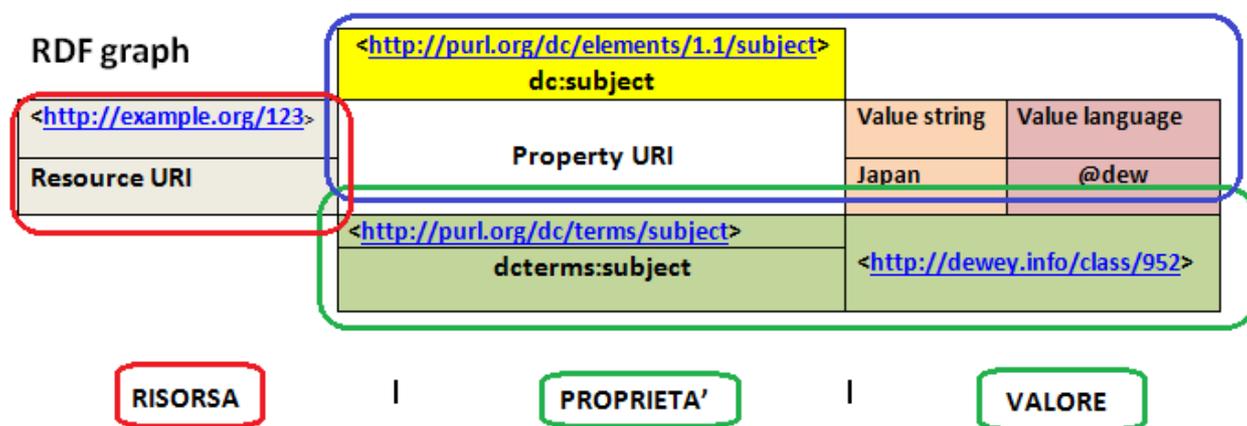


Figura 58. Esempificazione dei grafi RDF registrati in *Triple Store*

TripleStore è un database costruito appositamente e ottimizzato per la conservazione e il recupero delle informazioni (metadati) espresse tramite le triple RDF.

TripleStore possono essere visti come il vantaggio per la *performance* dei fornitori di dati (*Data Provider*), anche perché tutte le informazioni tracciate in un *TripleStore* possono essere recuperate attraverso un linguaggio di *Query* (es. il linguaggio di *Query del Fedora Software - Resource Index Query Service*; Figura 57).

¹ *Rendering* è una funzionalità del Software che prende le informazioni e visualizza il contenuto formattato sullo schermo.

Oltre alle Query, le triple RDF di solito possono essere importati/esportati al/dal sistema. Ad esempio, un valore "non letterale" (URI) che denota la proprietà del metadato "Soggetto" in relazione a una certa risorsa in un *Repository* - può essere importato/esportato al/dal da altri *Provider* che implementano *TripleStore* sui modelli di interrogazione dei grafi RDF nello spazio globale del Web dei Dati interconnessi.

Quindi, tramite *TripleStore* diverse applicazioni possono comunicare tra loro, creando i link reciproci per mezzo dei quali possono accettare o rifiutare certe asserzioni relative ai determinati metadati, recuperare o memorizzare in modo trasversale le rappresentazioni dei metadati, includere tutto o parte della descrizione dei metadati in riferimento ad un'altra rappresentazione, annotandola, o eseguendo le operazioni su di essa.

Le asserzioni di cui sopra si riferiscono a 4 o 5 stelle del già menzionato schema "*Tim's 5 star deployment scheme*" [V. sito]:

- (4) utilizzare gli URIs per indicare le cose (i valori, i fatti, le rappresentazioni), in modo che diversi agenti sul Web possano puntare tramite questi URIs agli specifici set di dati;
- (5) combinare tramite URIs i dati in modo sicuro con altri dati, nello schema globale di identificatori URIs, contribuendo allo sviluppo del contesto nei diversi domini informativi, alla ricchezza dei loro contenuti e allo scambio delle entità informative nello spazio globale di *Linked (Open) Data* (Web di Dati) del Web Semantico.

In ogni caso, la "triplicazione" di massa dei dati - senza alcun criterio di normalizzazione - nei *Repository* Istituzionali da parte degli *script* automatici deve essere evitata, in quanto tale non corrisponderebbe alle esigenze dello sviluppo delle rappresentazioni di qualità dei contenuti ospitati dai medesimi *Repository*, secondo Standard, Raccomandazioni e Linee guida. Una modellazione dei dati adeguata è un primo passo essenziale per contribuire alla qualità di qualsiasi applicazione.

Per quello che riguarda la creazione dei link automatici entro e tra diversi *TripleStore* RDF sul web, occorre usare i Vocabolari Standard per la descrizioni dei dati e metadati e dei loro valori [V. sito: *How to facilitate the linking between resources?*], altrimenti si corre il rischio di creare gli ammassi dei dati senza alcun criterio ovvero "RDF silos".

Un buon esempio di come possa essere raggiunto uno scambio dei contenuti e del contesto basati sui metadati sviluppati secondo le Raccomandazioni "LODE-BD" e gestiti in *TripleStore* RDF - è ben rappresentato dall'applicazione web "*OpenAGRIS_Beta*" collegato con la piattaforma di ricerca AGRIS (Figura 59). Qui il servizio "*OpenAGRIS_Beta*" genera i dati "*on the fly*" mappando le proprietà del metadato "Soggetto" con i valori "non-letterali" (URIs) del Thesaurus AGROVOC pubblicato sul Web come *Linked Open Data*. In questo modo la piattaforma AGRIS offre agli utenti la possibilità di ricerca e di recupero delle proprie risorse, a partire dai valori "*non-letterali*" (URIs) del Thesaurus AGROVOC. Inoltre, tramite questi valori le proprietà qualificate del metadato "Soggetto" del Repository AGRIS possono essere collegate in modo trasversale con gli stessi valori che qualificano le proprietà dei metadati in altri Repository, permettendo così agli utenti di recuperare anche le risorse correlate ospitate in altre piattaforme. Dalla piattaforma AGRIS è possibile, ad esempio, risalire alle risorse correlate ospitate nelle piattaforme "Nature.com" e "DBpedia" che utilizzano le rappresentazioni di *keywords* identificate dai valori URIs dell'AGROVOC.

Search AGRIS Find resources... Japan beta OpenAgris AGRIS About Feedback

Source:
Matica Srpska Library
 The Matica Srpska Library in its funds has more than 3,500,000 books and other publications.
 The Library receives the sample copy of all the publications of Serbia.

Biotechnology in Animal Husbandry (2011)
Integration of centralized biogas plant in cold-snowy region in Japan [2011]
 Umetsu, K., Obihiro University of Agriculture and Veterinary Medicine, Obihiro (Japan). Graduate School of Animal and Food Hygiene; Iwasaki, M., Obihiro University of Agriculture and Veterinary Medicine, Obihiro (Japan). Graduate School of Animal and Food Hygiene; Hokkaido Biomass Research and Development Company, Obihiro (Japan). Environmental Preservation Center; Ows, Obihiro (Japan).

Conference International Conference
Date of publication 2011
AGRIS Categories Renewable energy
AGROVOC English terms Japan; Biogas; Biogas production
AGROVOC French terms Japon; Biogaz; Biogaz production
AGROVOC Spanish terms Japon; Biogás; Biogás producción
Language English
Notes 8 ref.
Type Conference; Symposium
Resource - Is Part of <http://www.istoc.org/>
Resource - Is Part of <http://www.istoc.org/>
Journal Title **Biotechnology in animal husbandry (Journal)**
ISSN 1450-9156
Volume/Issue v. 27(3, book 1)
Abstract (English) A centralized biogas plant was built in Shikaoti town, Hokkaido, Japan to treat manure from 1320 cattle heads. The biogas plant was designed to operate at a feeding amount of 85.8 t/day, a hydraulic retention time (HRT) of 37 days and at a digester temperature of 38 deg C. In this study, the operational performance of biogas plant, utilization of digested slurry and economic balance were investigated. Since the working conditions of the plant became stable, the biogas production was 2,687 cubic m/day, 92% of produced biogas was consumed in power generation. Average methane concentration in produced biogas was 57.7%. The hydrogen sulfide (H2S) concentration was decreased to below 140 ppm as a result of bio-desulfurization and dry-desulfurization. The average power generated from that 54% of produced biogas was consumed in the facility operation, while 46% was sold to Power Company. Digested slurry were applied onto 602 ha grassland and agricultural fields. From the results of the operational performance and the economic balance evidence, it is evident that the centralized biogas plant has a positive effect on the local economy.

Agrovoc Keywords:
 • Electrical energy
 • economic situation
 • renewable energy
 • Biofuels
 • Biogas
 • Renewable resources
 • Japan

Biotechnology in animal husbandry (Journal)
 Once a database become a dataset of RDF triples, it can allow machine-to-machine querying
<http://aims.fao.org/standards/agrovoc/linked-open-data>
 Given the English preferred label, for example Japan, discover the URI and all the exact matches:

```
SELECT ?uri ?em {
  ?uri skos:prefLabel "Japan"@en .
  OPTIONAL { <http://aims.fao.org/aos/agrovoc/c_4039> skos:exactMatch ?em .
}
}

```

RESULT:
 uri em
 <http://aims.fao.org/aos/agrovoc/c_4039>
 <http://www.fao.org/countryprofiles/geoinfo/geopolitical/resource/Japan>
 <http://aims.fao.org/aos/agrovoc/c_4039> <http://eurovoc.europa.eu/1524>
 <http://aims.fao.org/aos/agrovoc/c_4039> <http://lod.nal.usda.gov/nalt/35135>

Integration of centralized biogas plant in cold-snowy region in Japan [2011]
 Shiroishi, K., Shikaoti (Japan). Environmental Preservation Center
 Takeuchi, Y., Hokkaido Biomass Research and Development Company, Obihiro (Japan)
 Oi, M., Shikaoti Public Office, Shikaoti (Japan). Environmental Preservation Center
 Kikuchi, S., Obihiro University of Agriculture and Veterinary Medicine, Obihiro (Japan). Graduate School of Animal and Food Hygiene
 Umetsu, T., Shikaoti Public Office, Shikaoti (Japan). Environmental Preservation Center
 Umetsu, K., Obihiro University of Agriculture and Veterinary Medicine, Obihiro (Japan). Graduate School of Animal and Food Hygiene
 Yasui, S., Community Think Tank Owls, Obihiro (Japan)
 Iwasaki, M., Obihiro University of Agriculture and Veterinary Medicine, Obihiro (Japan). Graduate School of Animal and Food Hygiene
 Ying, C., Obihiro University of Agriculture and Veterinary Medicine, Obihiro (Japan). Graduate School of Animal and Food Hygiene

the details of the record
bibliographic information and other information on the fly...
Resources from other datasets, linked through AGROVOC
Data from www.nature.com
 Policy: Public support for clean energy
 Spotlight on UK energy
 Biofuels need enforceable ethical standards
 Renewable fuels: Harnessing hydrogen
Data from DBpedia:
 Renewable energy
 Biogas
 Japan

Figura 59. Ricerca e recupero delle risorse nella piattaforma AGRIS tramite l'integrazione "on the fly" dei valori del Thesaurus *Linked Open Data* - AGROVOC

Le funzionalità di ricerca offerte dalla piattaforma "AGRIS" possono essere considerate uno dei migliori esempi che presentano tutti i componenti del sistema che lavorano insieme attraverso i valori del metadato "Soggetto" qualificati dal Thesaurus AGROVOC che utilizza gli URIs per identificare i concetti e la mappatura tra essi.

In corrispondenza di quanto descritto nel presente paragrafo, ne consegue che l'applicazione delle strategie di codifica e di implementazione dei metadati secondo le Raccomandazioni "LODE-BD" si basa sul concetto di "utilità per gli altri". Nel contesto di Archivi Istituzionali, l'utilità di "LODE-BD" può essere interpretata in termini di sviluppo e di implementazione di uno schema di metadati abilitati come *Linked Data* (appunto *Linked Data-enabled*). I metadati aventi queste caratteristiche possono essere riutilizzati da diversi agenti web, incluso le piattaforme di Archivi Istituzionali, contribuendo così al miglioramento trasversale della visibilità, dell'interoperabilità semantica e della comunicazione degli oggetti digitali su scala globale nel web.

5.5. Conclusioni

In questo Capitolo sono stati messi in luce diversi approcci e principi per la creazione di metadati estensibili e di qualità a servizio della descrizione dei contenuti di ricerca in un Archivio Istituzionale.

Secondo le "Linee Guida per la creazione e la gestione di metadati nei *Repository* Istituzionali" (CRUI, 2012) un Archivio Istituzionale di ricerca può ospitare le seguenti risorse digitali: Articoli di Riviste, *Paper* di Conferenze e Convegni, Poster di Conferenze, Patenti, Libri e loro parti, Annotazioni, Recensioni, Tesi di Dottorato, Tesi di Laurea, *Working Paper*.

Prima di passare alla fusione degli elementi di metadati - proposti dalle Linee Guida CRUI per la descrizione delle risorse digitali di cui sopra - con le strategie di codifica dei valori di metadati descritte entro le Raccomandazioni "LODE-BD" (*Linked Open Data – enabled Bibliographical Data*) (AIMS, 2012), nel Capitolo sono state passate in rassegna diverse questioni collegate logicamente al contesto della medesima fusione, cui l'obiettivo principale è quello di potenziare la *performance* dei metadati nella loro capacità di descrizione e di espressione tramite i valori "non-letterali" (*Uniform Resource Identifiers*, URIs) – elementi chiave per l'abilitazione dei metadati come *Linked Data* per l'esposizione dei contenuti sul Web Semantico.

Nello specifico, il Capitolo ha affrontato le seguenti tematiche:

1) L'importanza della pubblicazione delle entità informative con l'ausilio della tecnologia *Linked Data*, un candidato molto promettente e altamente raccomandato dalla comunità Web per affrontare una delle più grandi sfide nel settore della gestione delle informazioni intelligenti: l'uso del Web come piattaforma per la connessione e l'integrazione di una considerevole quantità e diversità tipologica di *datasets* (risorse digitali e loro parti costitutive) che risiedono sopra il livello di documenti ipertestuali del Web tradizionale.

Per esprimere le entità e le connessioni come *Linked Data*, si ricorre alle connessioni basate sul modello dei dati strutturati *Resource Description Framework* (RDF) e sugli identificatori web (tramite il protocollo HTTP) URIs.

Publicare i *dataset* informativi come *Linked Data* non significa solo pubblicarli in modo innovativo sul Web. *Linked Data* - in primo luogo – significa creare le relazioni, in modo che persone e sistemi web possano esplorare diverse reti di conoscenza create sui dati intercorrelati. Con *Linked Open Data*, quando si dispone di alcuni dati Aperti (pubblicati tramite Licenze e Formati aperti) di partenza e si può così esplorare quello spazio di dati interoperabili del Web, conosciuto come *Linked Data* o Web 3.0, che è espressione pratica del Web Semantico.

2) L'importanza dell'uso dei Sistemi di Organizzazione della Conoscenza (*Knowledge Organization Systems* o KOS) che comprendono vocabolari controllati e schemi semantici con cui valori è possibile qualificare le proprietà dei metadati. KOS offrono un modo sistematico per organizzare, normalizzare, arricchire semanticamente, accedere e recuperare i contenuti delle risorse informative, attraverso la proposta d'uso di termini predefiniti (voci di autorità), di sinonimi, varianti ortografiche, gerarchie tesaurali, strumenti concettuali analoghi (classi, associazioni).

Attualmente molti Vocabolari Controllati e Schemi Semantici (Classificazione, Ontologie) sono disponibili *online* e i loro valori sono identificati con gli indirizzi web <http://URIs>, quindi con i valori "non letterali" mappati con i corrispondenti valori letterali (stringe di Soggetto, Classi, voci di Thesauri).

L'affiancamento con tali valori delle proprietà dei metadati d'Archivio non solo rappresenta un approccio innovativo per il controllo di autorità e per l'arricchimento semantico dei metadati, ma anche per lo sviluppo degli "accessi multifocali" ovvero degli accessi unici a diverse risorse distribuite in modo trasversale in diversi sistemi sul web.

3) Alcuni approcci, documenti e principi per la creazione di metadati estensibili di qualità al servizio della descrizione dei contenuti di ricerca. Viene fatto riferimento ai sei principi dei

cosiddetti "Good metadata" definiti da NISO (*National Information Standards Organization*), secondo cui "Good metadata":

- a) sono conformi agli Standard comunitari, cioè riconosciuti e accettati internazionalmente per la descrizione delle risorse e delle loro collezioni, per soddisfare le esigenze informative degli utenti (attuali, futuri e potenziali), e le esigenze locali dei sistemi per la rappresentazione, ricerca e il recupero dei contenuti digitali;
- b) supportano l'interoperabilità tra le risorse che descrivono, entro e al di fuori dei sistemi che esse ospitano;
- c) si appoggiano a strumenti del controllo di Autorità, Standard e Schemi semantici per descrivere gli oggetti in modo normalizzato, dal punto di vista semantico e per creare i collegamenti tra gli oggetti affini nel contenuto e contesto;
- d) contengono chiare dichiarazioni sulle condizioni e termini d'uso per ogni oggetto digitale;
- f) sostengono la conservazione delle risorse (incluso loro oggetti costitutivi e metadati) a lungo termine.
- g) sono considerati come oggetti di contenuti (*Content Objects*) aventi proprietà, valori e requisiti di oggetti di qualità, in possesso delle caratteristiche di identificazione univoca, di autorità, autenticità, e persistenza a lungo termine.

"Good metadata" richiedono la comprensione sia dei dati (contenuti) - che stanno per descrivere - che degli Standard sui quali le medesime descrizioni possono essere appoggiate.

Nella sezione "Validazione di metadati" delle "Linee Guida per la creazione e la gestione di metadati nei Repository Istituzionali" [V. sito, pp. 11-12] è stato sottolineato che la qualità dei metadati, a sua volta, determina la qualità delle funzioni svolte e dei servizi offerti sia dai Repository/Archivi (*Data Provider*, fornitori di dati) che dagli aggregatori (*Service Provider*, fornitori di servizi) che raccolgono i dati da *Data Provider* via il Protocollo OAI-PMH, contribuendo a migliorare il contesto di interoperabilità all'interno delle piattaforme conformi a *Open Archive Initiative*, tramite la condivisione comune delle mappature tra diversi schemi di metadati.

4) Strategie delle Raccomandazioni "LODE-BD" (*Linked Open Data – enabled Bibliographical Data*) che vengono rispecchiate principalmente negli Alberi decisionali "Decisions Trees" proponenti i passi ragionati per:

- a) la scelta degli elementi di metadati al supporto descrittivo delle risorse nei Repository;
- b) la qualificazione delle proprietà dei metadati tramite diversi valori – per la loro identificazione, l'autenticazione, la normalizzazione e la semanticizzazione, invitando i gestori dei Repository da privilegiare nella scelta dei valori "non-letterali" identificati dagli indirizzi web <http://URIs> e perciò capaci di abilitare i metadati – memorizzati nei Software RDF *Triplestores* - a diventare *Linked Data*.

5) Alcuni vocabolari controllati per l'abilitazione delle proprietà di metadati "RightsHolder", "Licence", "AccessRights", "Rights" come *Linked Data*.

6) Mappatura dei dati bibliografici (metadati) descrittivi le risorse nei Repository con i marcatori dello "Schema.org", per una migliore comunicazione delle medesime risorse sul Web.

La potenza dei marcatori di un metalinguaggio web flessibile XML - con il quale vengono espresse diverse risorse sul Web - è senz'altro un grosso vantaggio, perché così si evita la rigidità di altri linguaggi web, come l'HTML, attribuendo dei <tag> o delle etichette *customizzate* (personalizzate) ad una varietà di attributi delle risorse aventi un proprio significato. Dall'altra parte, tanta flessibilità può produrre diverse versioni di Ontologie (dizionari) esterne simili (es. diverse versioni di metadati RDF), minando la standardizzazione delle informazioni. È per questo motivo che i motori di ricerca Google, Yahoo! e Bing - per operare in un'ottica di standardizzazione delle meta-informazioni sui contenuti delle pagine, e per dare a chi gestisce i

siti web un'unica soluzione da utilizzare - hanno deciso di comune accordo di sostenere il progetto "Schema.org".

"Schema.org" è un'ontologia con la quale è possibile spiegare e dare informazioni aggiuntive su diverse tipologie di contenuto: eventi, libri, prodotti, persone, musica, offerte e recensioni. "Schema.org" potrebbe essere anche un buon incentivo per i metadati delle risorse degli Archivi Istituzionali affinché i motori di ricerca possano reperire da essi più informazioni possibili e riconoscere dettagli delle descrizioni altrimenti difficili da identificare e isolare sul web.

7) Alla fine è stato proposto un Profilo specifico di metadati basato sulle Linee Guida CRUI e sulle Raccomandazioni LODE-BD per la descrizione delle risorse di ricerca negli Archivi Istituzionali.

Questo profilo è allocato all'interno di uno schema di metadati rappresentati dai termini ben definiti, affiancati dalle specifiche di compilazione, dalle strategie dell'allineamento (mappatura) ai termini di metadati più specifici, dai valori di Vocabolari Controllati e Schemi semantici (dove possibile e necessario) che permettono ai metadati di base di diventare:

- più efficienti nel contesto di controllo d'autorità,
- più ricchi nei loro profili semantici e
- più accessibili e fruibili sul Web per mezzo della scelta dei valori "non-letterali" abilitanti i metadati a diventare *Linked Data* nel contesto del Web Semantico.

Sviluppare e implementare (secondo le strategie descritte entro le Raccomandazioni LODE-BD "Design-time" e "Run-time") schemi di metadati allineati in tutto o in parte secondo il paradigma di *Linked Data* arrecherà diversi vantaggi alle risorse degli Archivi Istituzionali e alle loro collezioni, entro e fuori dei medesimi Archivi, in quanto tale approccio proietta gli effetti positivi su:

- lo scambio delle informazioni tra i metadati entro e fuori dell'Archivio di riferimento;
- la normalizzazione trasversale dei valori metadati, ovvero nel modo incrociato tra diversi sistemi che impiegano per l'indicizzazione delle proprie risorse gli stessi Vocabolari controllati disponibili ed identificabili con i loro valori online tramite gli indirizzi web URIs;
- l'espressività delle relazioni semantiche impiegate nei processi di navigazione e di ricerca delle risorse - intese come contenuti (*Content Objects*) correlati entro e tra sistemi basandosi sui valori web "non-letterali" <http://URIs> scelti in conformità con i passi degli Alberi Decisionali e delle strategie di codifica contenute nelle Raccomandazioni "LODE-BD";
- l'integrazione e la partecipazione delle risorse degli Archivi Istituzionali nella nuvola dei *dataset* collegati a livello globale in rete (Web) - *Linked Open Data Cloud*;
- il contributo degli Archivi Istituzionali allo spazio aperto dei contenuti di ricerca collegati - *Open Research Commons Space*.

La sfida per il futuro della ricerca è di creare una forte interconnessione e uno scambio di informazioni (scientifiche, finanziarie, amministrative, di supporto alla valutazione) tra tutti gli attori del sistema Universitario internazionale, definendo Linee guida e Standard, nonché gli Schemi di metadati descrittivi con gli elementi abilitati come *Linked Open Data* riconosciuti e apertamente interscambiabili tra diversi agenti in rete, dopo aver identificato esigenze comuni e specificità.

5.5. Conclusions

This Chapter highlighted different approaches and principles for the creation of extensible and qualitative metadata for the description of research contents in Institutional Repositories (IRs). According to the "Guidelines for metadata creation and management in Institutional Repositories" (CRUI, 2012), an IR can accommodate the deposit of the following digital resources: Journals articles, Papers and Posters of Conferences, Patents, Books and their parts, Reviews, Master and Ph.D. Thesis, Working Papers and other materials having research value. Before adapting metadata elements - proposed by the "Guidelines CRUI" to describe the above mentioned digital resources – to encoding strategies provided by "LODE-BD (Linked Open Data - enabled Bibliographical date) Recommendations" (AIMS, 2012), several issues connected to metadata quality were reviewed, as well as to the performance of their values as Linked Data – enabled units. This is possible through qualifying metadata values with "non-literal" values expressed by *Uniform Resource Identifiers (URIs)* based on web <http://> addresses, which should be allocated within RDF structures.

Particularly, the Chapter addressed the following issues:

1. The importance of publication of information entities by means of Linked Data technology, which is a highly recommended and promising candidate to address one of the biggest challenges in the field of intelligent information management: the use of the Web as a platform connecting and integrating a considerable quantity of digital datasets (information resources and their constituent parts) residing above the level of hypertext documents of the traditional Web.

To publish information datasets as Linked Data is not just to publish them in an innovative way on the Web. First of all, Linked Data means creating connections, so that people can explore different web-based contents and cross-related knowledge networks. Working with Open datasets (published with Open Licenses and formats), different related datasets can be created and freely accessed, favoring easy exploration of the global space of Linked Open Data, a considerable practical expression of the Semantic Web.

2. The importance of implementation of Knowledge Organization Systems (KOSs) including controlled vocabularies and semantic schemes (such as Thesaurus, Classifications, Metadata Schemas, Ontologies), whose values can be chosen to qualify metadata properties. KOSs offer a systematic approach to organize, normalize, enrich semantically, access and retrieve contents of information resources. It is possible via predefined terms (authority voices), synonyms, spelling variants, thesaurus hierarchies, similar conceptual tools (classes, associations). Currently different KOSs are available on the Web as Linked (Open) Data. The values of these KOSs are identified with "non-literal" values referred to corresponding "literal" (Subject strings, classes, Thesaurus items) units. Qualifying metadata properties with such types of values not only offers an innovative approach to Authority Control and semantic enrichment of metadata properties, but also contributes to the creation of unique and multipurpose accesses to different correlated resources distributed across the Web.
3. Some existing approaches, documents and principles for creating extensible and qualitative metadata describing digital research contents. In this context six principles of so-called "Good metadata" defined by NISO (National Information Standards Organization) were mentioned. According to NISO, "Good metadata" is:
 - a) complying with widely shared community Standards recognized and accepted internationally to describe resources and their collections and to meet information needs of users (current, future, potential). "Good metadata" is also as a powerful tool to accommodate local requirements and needs of digital informational systems concerning formation, acquisition, registration, classification, preservation, representation/visualization, search and retrieval of digital contents;
 - b) enabling interoperability among resources within and outside systems;

- a) referring to Authority control tools, as well to semantic schemes, whose units may describe objects both in a normalized way and from the semantic point of view. Semantic and normalized metadata elements are suitable to create links among information objects having elements (referred to content and context) in common;
- b) containing clear statements on the conditions and terms of use for each digital object;
- c) supporting digital preservation of resources in long term perspective;
- d) being considered as "Content Objects" itself having its own properties, values and requirements of quality information items based on the unique persistent identifiers (PID) sustaining long-term access to information resources. The values of "Good metadata" are subjected to Authority Control, thus proving high sustainability and authenticity of objects they describe.

"Good metadata" requires an understanding of both data (Content Objects) it is going to describe, as well as Standards that can support these descriptions.

In the section "Validation of metadata" of the "Guidelines for metadata creation and management in Institutional Repositories" [V. website, pp. 11-12] it is stressed that metadata quality, in turn, determines the quality of functionalities and services offered by Data Providers (Repositories or Archives of digital contents) and Service Providers. These last are aggregating (collecting, harvesting) metadata from Data Providers via OAI-PMH (Open Archive Initiative Protocol for Metadata Harvesting) in order to create services with added value.

4. "LODE-BD" (Linked Open Data - enabled Bibliographical Data) strategies being promoted by AIMS web Community as an important tool guiding the choice and encoding of metadata properties according to local needs and choice of "literal" and "non-literal" values. These strategies are reflected in "LODE-BD" Decision Trees supported by explanation Tables providing logical steps to:

- a) choose metadata elements describing information resources;
- b) qualify metadata properties by appropriate values. This step is very important for identification, normalization, authentication, standardization and semantization of metadata properties. Encoding metadata values (stored in RDF Triplestores) with "non-literal" enables metadata to be expressed (published) and grasped on the Web as Linked (Open) Data.

5. "LODE-BD" Decision Tree referring to metadata elements describing Intellectual Property issues ("RightsHolder", "Licence", "AccessRights", "Rights"), extended with logical steps guiding choice of Licenses in order to identify a correct use of information resources. A Table with some considerable controlled vocabularies enabling metadata properties to acquire characteristics of Linked Data entities is presented.

6. Mapping of bibliographic data (descriptive metadata) with mark-ups of "Schema.org" web Ontology for its better interpretation and retrieval on the Web. The power of mark-ups provided by flexible XML web meta-language syntax is certainly of a considerable importance to avoid non-flexibility of other web languages such as, for example, HTML. On the other hand, the flexibility offered by XML may produce different versions of the same web Ontologies (e.g. different versions of RDF structured metadata), undermining the standardization of information. For this reason and in order to operate with meta-information of web pages in a standardization way, the search engines such as Google, Yahoo! and Bing have commonly decided to support the "Schema.org." Project. "Schema.org" is a web Ontology through which it is possible to explain and to give additional information about different types of digital content, referring, for example, to events, e-books, journals, products, people, music, reviews and other information entities. "Schema.org" is a suitable tool providing a good incentive to sustain expressivity of metadata assets in Institutional Repositories so that search engines can find and retrieve these assets as easier as possible, recognizing details of metadata descriptions, otherwise to be difficult identified and isolated on the Web.

7. A specific metadata profile based on the “Guidelines CRUI” and “LODE-BD Recommendations”. This profile is allocated within a metadata schema presented by well-defined metadata terms supported by detailed specifications for compilation and alignment (mapping) strategies in terms of cross-referring to more specific metadata, to values of some controlled vocabularies and semantic schemes (Thesaurus, Ontologies, Classifications), where it is possible and necessary. Such an approach allows basic metadata to become:

- a) more efficient in the context of Authority Control;
- b) richer in their descriptive profiles and semantics;
- c) more accessible and usable on the Web through qualifying of its properties by "non-literal" values enabling metadata to become Linked Data in the context of Semantic Web.

Encoding metadata elements in compliance with LODE-BD “Decision Trees” and implementing them according to the LODE-BD "Design-time" and "Run-time" strategies will bring several advantages to IRs resources, such as:

- better exchange of information among metadata elements based on connections that can be created within and outside IRs technological confines;
- cross-way normalization of metadata values in different Web information management systems hosting a variety of types of information resources and recurring to the same “non literal” (<http://URIs>) values of Linked Data-enabling vocabularies. This presents a considerable potential for harmonization of interaction among a variety of digital contents within and outside different IRs platforms on the web;
- integration and participation of IRs resources in the Linked Open Data Cloud viewing the aggregation of different information entities and their connections published with help of Linked Data technology, in order that these entities are understood, shared and enriched in the global Semantic Web space of interconnected data - Web of Data (Web 3.0);
- contribution of IRs contents to the so called Ecosystem - Open Research Commons Web Space. Such a framework should promote and sustain publishing, describing, representing, linking, reusability and interoperability of research contents in innovative way, with respect to Open Standards fostering open cross-collaboration and fertilization of contents among different research communities operating on the web.
- The challenge of future research space on the web is to create strong interconnections and provide rich exchange of information (scientific, financial, administrative, didactic, evaluating) among all institutions of the international University system. This is possible through conforming to common Open Standards, Guidelines and Recommendations (e.g. such as “LODE-BD”), as well as using widely accepted and shared metadata schemas and their mappings (cross-walks), as well as widely-accepted metadata application profiles, controlled vocabularies and semantic schemes enabling basic metadata elements to acquire characteristics of Linked Open Data entities. Datasets published and exposed on the Web as Linked Data through Open Licenses and formats are high promising candidates to:
 - a) enhance interchange of digital contents among different agents acting on the Web;
 - b) improve and enhance access to correlated digital contents;
 - c) supplement, in a transversal way, search and navigation services with added value, getting closer to user information needs, focused primarily on the necessity to retrieve qualitative, and semantically correlated in a creative way, information contents.

Conclusioni della Tesi.

Comunicazione dell'informazione scientifica: le nuove frontiere

"l'informazione è una differenza che fa differenza" [Bateson, 1989]

L'obiettivo del presente lavoro di Tesi è stato quello di entrare in *medias res* delle modalità di promozione di un nuovo modo di veicolare la conoscenza scientifica tramite il movimento internazionale *Open Access*. La disseminazione dei lavori di ricerca per mezzo di questo movimento si basa su due assi portanti ovvero due canali comunicativi appoggiati sulla rete Internet:

- 1) Riviste *Open Access* (la strategia *Gold Road*);
- 2) Archivi *Open Access* (la strategia *Green Road*).

Le medesime strategie sono considerate complementari (pubblicazione, deposito, consultazione) e corrispondono pienamente alla necessità di:

- a) potenziamento delle modalità della comunicazione dell'informazione scientifica finanziata con i fondi pubblici che - per la definizione stessa - deve essere di dominio pubblico cioè "a portata di tutti", appunto, "*Open Access*". Questo obiettivo può essere raggiunto rendendo esplicitamente il materiale di ricerca e i suoi dati descrittivi (dati bibliografici, metadati) di Pubblico Dominio tramite, ad esempio, licenze PDDL (*Public Domain Data Licence*) o CC0 (*Creative Commons Zero*);
- b) creazione dei canali digitali di comunicazione attiva e aperta tra la comunità di ricerca e una varietà di comunità di utenti.

In riferimento alla **PARTE I**, il **Capitolo 1** ha ricostruito i presupposti della nascita, le caratteristiche e le principali Dichiarazioni, così come le Politiche e le iniziative di portata internazionale del movimento *Open Access*.

Nel **Capitolo 2** l'attenzione è stata focalizzata sulla seconda strategia dell'*Open Access* chiamata *Green Road* riferita al dominio di Archivi *Open Access* implementati sulle piattaforme (*Software*) aperte in quanto sono conformi ai requisiti dell'*Open Archive Initiative* (OAI), in particolare al protocollo dell'interoperabilità (interazione) tecnica tra i metadati - *Open Archive Initiative Protocol for Metadata Harvesting* (OAI-PMH).

L'indagine del campo teorico riferito agli Archivi aperti *Open Access* è stata ristretta all'osservazione degli Archivi Istituzionali (*Institutional Repositories, IR*) - definiti come luoghi digitali con una chiara *identità istituzionale* (in riferimento a Università ed Enti di ricerca) per la raccolta o il deposito (auto-archiviazione), la consultazione e la disseminazione del materiale di ricerca prodotto dai membri delle Istituzioni di riferimento. Una chiara identità istituzionale (es. l'*Archivio Istituzionale dell'Università di Vienna*) è la caratteristica che rende diversa questa tipologia di Archivi dagli Archivi Disciplinari (*Subject-based Repository*) *Open Access*.

Dopo la descrizione degli scopi e delle funzionalità principali degli Archivi aperti Istituzionali, ci si è inoltrati nel campo degli strumenti e delle procedure che possono essere coinvolti nelle questioni riguardanti la loro programmazione, etica e qualità. In particolare, gli Archivi aperti Istituzionali sono stati osservati dall'angolatura:

- 1) delle Linee di indirizzo (Linee Guida, Raccomandazioni) a supporto della loro implementazione e gestione, cui si dovrebbe far riferimento nell'allestimento delle eventuali Politiche di Archivi;
- 2) dei loro possibili collegamenti con *Current Research Information Systems* (CRIS). In questa prospettiva è stato focalizzato un potenziale compito degli Archivi Istituzionali - particolarmente quelli integrati con i servizi per la *Governance* dei dati di ricerca sviluppati in Italia: "*SurPlus Open Archive*" e "*U-GOV Ricerca*" - nel contribuire al processo della valutazione della ricerca a livello istituzionale e nazionale;
- 3) dell'etica riferita al rispetto dei contratti editoriali e della proprietà intellettuale durante le procedure di deposito e di consultazione dei prodotti di ricerca;

4) dell'importanza della definizione delle Politiche per la programmazione della conservazione dei contenuti digitali a lungo termine, secondo il modello dello Standard ISO:14721 OAIS (*Open Archival Information System*);

5) della necessità di stesura – prendendo lo spunto dalle Politiche dell'Archivio - di un *Data Management and Process Plan* il quale deve contenere la descrizione di tutte le procedure e responsabilità necessarie per la gestione di qualità dell'intero ciclo di vita dei contenuti digitali, in prospettiva di medio-lungo termine e per una migliore consapevolezza degli obiettivi e dei rischi impliciti nell'organizzazione di tutti gli assetti dell'Archivio.

Prestare una dovuta attenzione agli aspetti appena elencati da parte dei programmatori e gestori degli Archivi Istituzionali è un compito di rilevante importanza per contribuire alla *qualità di comunicazione dell'informazione scientifica* tramite i canali di rete.

La *qualità di comunicazione dell'informazione* degli Archivi Istituzionali non può prescindere da un'adeguata impostazione:

- della metadazione standardizzata delle loro risorse;
- dai metodi che forniscono l'*accesso* per "Soggetto" alle risorse.

Le proprietà di questo tipo d'accesso devono essere qualificate e arricchite dai valori di Vocabolari Controllati e Schemi semantici (es. Terminologie, Thesauri, Classificazioni).

L'*accesso semantico* ai contributi depositati negli Archivi deve essere inoltre armonizzato con gli indici delle bibliografie e dei Cataloghi digitali delle Biblioteche così come con gli indici di risorse di ogni altra tipologia, per dare luogo al cosiddetto "*accesso multifocale*" ai contenuti, ovvero il *medesimo accesso semantico* a tutte le tipologie di risorsa, appianando il concetto di differenza tra i metadati dei vari silos che contengono informazioni e testi;

- dell'*accesso permanente* alle *copie autentiche* delle risorse depositate che devono essere basate sugli identificatori unici e persistenti (duraturi).

In prospettiva della garanzia della qualità dell'accesso permanente alle proprie risorse informative, gli Archivi Istituzionali possono assumere la veste di prestazioni dei *Trusted Digital Repositories* ovvero degli Archivi digitali implementati e gestiti in vista delle *Politiche della conservazione dei contenuti digitali di lunga durata*. Queste Politiche devono riflettere tutte le azioni e gli interventi necessari per risolvere il problema delle versioni, degli identificatori, delle migrazioni e dei *back-up* dei file, per garantire un costante accesso agli oggetti digitali che conservano le caratteristiche di affidabilità, autenticità e integrità per tutto l'arco di tempo in cui vengono considerati utili.

Mentre lo Standard OAIS (*Open Archival Information System*) indica col termine "*Archive*" (Archivio) un tipo di organizzazione il cui scopo è quello di conservare l'informazione (pacchetti di oggetti informativi e di metadati a loro supporto) in modo che questa sia accessibile e fruibile a lungo termine da parte di una determinata comunità, i requisiti di *Trusted Digital Repositories* (ISO 16363:2012 - Audit and Certification of Trustworthy Digital Repositories) si riferiscono al campo semantico "*Archive*" tramite un termine composto – "*Digital Repository*" [Deposito Digitale].

Dal punto di vista di un Archivio (*Repository*) aperto Istituzionale, ambedue i termini – "*Archive*" e "*Digital Repository*" sono spesso utilizzati in modo interscambiabile. Da ciò deriva che un Archivio (*Repository*) Istituzionale deve essere considerato sia come un luogo per il "*Deposito Digitale*" delle risorse informative sia come un "*Archivio*" in cui vengono integrati tutti gli strumenti necessari per la conservazione a lungo termine delle medesime risorse, portando come esempio – l'Archivio Disciplinare *Open Access* - ArXiv, cui contenuti sono in rete da decenni e sono ancora consultabili. Intanto, non si devono confondere Depositi Digitali e Archivi:

- né con le Biblioteche Digitali, che raccolgono le informazioni digitali e ne permettono l'accesso, ma senza necessariamente impegnarsi a conservarle a lungo termine,
- né con gli Archivi di Dati, che mirano a una conservazione a lungo termine, ma i cui contenuti sono solamente insiemi di dati statistici.

L'implementazione di Archivi digitali in vista di *Trusted Digital Repositories* ben documentati con Politiche, procedure e pratiche esaurienti ed esplicite - implica la possibilità di *certificazione* (valutazione sia interna che esterna) di loro qualità sul *Programma di conservazione digitale*. La prospettiva di certificazione deve stimolare:

- la disseminazione di *Best Practices*, promozione di Raccomandazioni e Linee Guida col riferimento alla conservazione dei contenuti a lungo termine (es. "Linee Guida per gli Archivi Istituzionali" - sessione "*Preservazione*". Gruppo *Open Access* della Commissione Biblioteche CRUI);
- lo sviluppo di *sinergie tra ricerche* sulla conservazione del digitale nei settori di Archivi Istituzionali e *Trusted Digital Repositories*;
- le compagnie ICT (*Information Communication Technology*) e gli sviluppatori di *Software* per le piattaforme di Archivi aperti per approfondire problematiche tecniche legate alla conservazione del digitale.

In riferimento alla **PARTE II**, nel **Capitolo 3** è stata eseguita un'indagine su un campione delle piattaforme (nazionali e internazionali) digitali a sostegno della comunicazione del sapere e del Libero Accesso ai contenuti di ricerca in rete. Il campionamento dell'indagine ha riguardato 23 infrastrutture compatibili col protocollo dell'interoperabilità - *Open Archive Initiative Protocol for Metadata Harvesting (OAI-PMH)* e ha incluso gli Archivi (Istituzionali e Disciplinari) *Open Access* del materiale col valore scientifico e storico-culturale e il Portale di ricerca "*The European Library*" delle risorse aggregate da 47 biblioteche europee. La maggior parte degli Archivi rappresentanti l'oggetto d'indagine sono stati valutati dal servizio online *Ranking Web of Repositories* (pubblicato dal 2008) in quanto al livello di posizionamento "tra i pari".

L'obiettivo dell'indagine campionaria è stato quello di individuare:

- i requisiti tecnici delle piattaforme di riferimento (Software; dichiarazioni sui metadati e sull'accesso semantico; Statistiche d'uso dei materiali; meccanismi a supporto dell'accesso a lungo termine e di ricerca/navigazione; programmazione dei livelli d'accesso);
- le funzionalità aggiuntive/personalizzate di rappresentazione e ricerca dei contenuti,
- i meccanismi che connettono le piattaforme con altre infrastrutture digitali di ricerca e di comunicazione (es. con i canali di *Social Networks*).

È stato rilevato che tutte le infrastrutture sviluppate sulle stesse piattaforme di *Software* presentano caratteristiche simili, in quanto riflettono le funzionalità di *default* dei medesimi *Software*. Ciononostante, diverse tra queste piattaforme introducono le caratteristiche peculiari, personalizzando e amplificando così le funzionalità che riguardano il *browsing*, la ricerca e la presentazione/visualizzazione dei contenuti.

Tra i buoni esempi di piattaforme con servizi di ricerca e *browsing* personalizzati vi sono quelli in cui le proprietà di metadati descrittivi (riferiti al Nome dell'Autore, Ente, Soggetto, Luogo) le risorse depositate sono affiancate dagli strumenti di autorità come *Vocabolari controllati* e *Schemi semantici di Classificazione*. L'uso di questi strumenti negli Archivi Istituzionali potrebbe condurre alle nuove frontiere della ricerca sulla comunicazione dell'informazione scientifica, specialmente per quanto riguarda:

1) l'*implementazione* negli Archivi di *Vocabolari controllati* e *Schemi semantici* pubblicati online come *Linked Open Data (LOD)* con i "valori non - letterali" cioè rappresentati dagli indirizzi Web (basati sul protocollo web *http*) URIs. Questo approccio permetterà di creare:

a) un *confronto innovativo* fra *parole* contenute nei documenti (indici estratti sulla base della loro *occorrenza*) o scelte per gli *Abstract* - e valori della terminologia identificata con gli indirizzi Web URIs, per recuperare:

a.1) sia le descrizioni normalizzate e semantiche dai Thesaurus e Schemi di Classificazione attribuendo la garanzia formale dell'espressione agli indici descrittivi del materiale d'Archivio,

- a.2)** che le risorse che derivano dall'uso della medesima descrizione entro e fuori l'Archivio di riferimento, come il risultato del succitato “*accesso multifocale*”. basato sui collegamenti semantici appoggiati sui “*valori non – letterali*” rappresentati dagli indirizzi web URIs.

È chiaro che l'*indicizzazione semantica* dei documenti negli Archivi Istituzionali rappresenta uno degli aspetti più importanti per potenziare l'organizzazione e l'accesso ai medesimi documenti e alla conoscenza che essi veicolano, ma potrebbe rappresentare anche l'aspetto più complesso della loro organizzazione, in quanto l'impiego dei vari Thesauri, Classificazioni o Terminologie strutturate (anche se esse rappresentano la stessa area) potrebbe causare forti disaccordi sul piano linguistico, strutturale e semantico, con divergenze particolarmente gravi come, ad esempio, tra le Classificazioni Disciplinari (es. lo schema di classificazione JITA per l'ambito LIS/*Library Information Science*) e le Classificazioni generali (es. la Classificazione Decimale Dewey, la Classificazione Decimale Universale, o la Classificazione della Library of Congress). L'impiego dei *Vocabolari controllati e Schemi semantici* con i “*valori non – letterali*” [V. siti: CKAN Data Hub, *Linked Open Data Cloud*, Rapporto del Gruppo *Library Linked Data - “Datasets, Value Vocabularies, and Metadata Element Sets”, Linked Open Vocabularies (LOV)*] privilegiando quelli in cui sono previste le mappature con i valori di altre Liste semantiche e di autorità - è una risposta per superare i succitati limiti.

2) La necessità di *sviluppo* dei *Thesaurus specializzati* per gli Archivi Istituzionali, collegando (anche tramite le equivalenze linguistiche) le voci di questi Thesaurus con quelli ampiamente diffusi e condivisi a livello nazionale e internazionale e disponibili online come *Linked Open Data*.

Questo approccio permetterà di:

- a) potenziare il linguaggio descrittivo delle risorse secondo le esigenze locali dell'Archivio e nello stesso tempo fare il riferimento a un lessico semantico ampiamente riconosciuto e condiviso, identificabile e recuperabile in rete tramite gli identificatori URIs;
- b) creare un lessico interoperabile tra Archivi diversi;
- c) amplificare la rete di collegamenti entro una ricca gamma di risorse in rete a livello globale, favorendo così diverse forme di cooperazione e condivisione nel reperimento efficace ed esaustivo delle risorse informative.

Tra i fattori catalizzanti il valore e il successo degli Archivi Istituzionali sono stati individuati:

- 1) la necessità di seguire un piano nazionale/europeo/internazionale per un futuro organico di sviluppo e di gestione degli Archivi, come, ad esempio, proposto dalle Linee Guida CRUI, *DRIVER*, *OpenAIRE* [V. “Sitografia”];
- 2) l'importanza dell'allestimento degli Archivi Istituzionali con gli schemi di metadati e loro mappature ampiamente riconosciuti e condivisi a livello nazionale/europeo/internazionale, per favorire una migliore comunicazione e integrazione tra i contenuti di diversi Archivi e sistemi online, compreso sistemi informativi per la ricerca CRIS (*Current Research Information Systems*);
- 3) l'integrazione degli Archivi Istituzionali nelle questioni amministrative legate all'Anagrafe della ricerca;
- 4) l'affiancamento agli elementi di metadati dai meccanismi che li possono abilitare come *Linked Data*, secondo le strategie di codifica *LODE-BD (Linked Open Data (LOD)-enabled bibliographical data)*, le quali guidano i metadati e le risorse che essi descrivono da database locali nell'universo di *LOD (Linked Open Data)*, che contiene una diversità di *dataset* (pacchetti di contenuti) pubblicati in rete dai vari soggetti produttori e interconnessi in base alle proprietà comuni;
- 5) l'implementazione di *identificatori unici persistenti*, e il supporto dei *formati* per la *conservazione a lungo termine* dei contenuti digitali;

- 6) l'implementazione di servizi a valore aggiunto per gli aggiornamenti e la comunicazione attiva con l'utenza (es. sezioni come "Statistiche d'uso", "Commenti critici collaborativi", "Annotazioni", *Forum, Blog*),
- 7) lo sviluppo dei collegamenti tra diversi Archivi di ricerca e altri Portali di ricerca a livello globale in rete, investendo nello sviluppo dello spazio di ricerca virtuale comune "*Global Scientific Data Space*".

In riferimento alla **PARTE II**, nel **Capitolo 4** è stato presentato un *ToolKit* basato sullo stato dell'arte di alcuni aspetti a supporto del "Profilo Tecnico" di Archivi digitali di ricerca. In particolare, sono stati riportati i riferimenti a rilevanti osservazioni teoriche, Progetti, strumenti, illustrazioni di metodologie e soluzioni pratiche in collegamento alle seguenti sezioni:

- 1) Panorama delle piattaforme;
- 2) Metadati a supporto dei contenuti digitali;
- 3) Controllo di autorità dei valori di metadati;
- 4) Associazione di metadati con valori semantici;
- 5) Conservazione a lungo termine;
- 6) Digitalizzazione;
- 7) Scoperta delle risorse: paradigmi di ricerca e navigazione.

I contenuti con cui è stato allestito il Toolkit non hanno avuto la pretesa di essere esaustivi. Piuttosto, possono essere considerati come i punti da cui partire per elaborare un'ulteriore strategia per il potenziamento del "Profilo Tecnico" degli Archivi Istituzionali, attraverso i passi da eseguire in maniera incrementale per il raggiungimento di livelli di sviluppo successivi o in maniera di riadattamento alle esigenze di funzionalità aggiuntive ed identità funzionali richieste dal progetto dinamico (in continua implementazione) di sviluppo delle funzionalità di un Archivio Istituzionale.

In relazione alla **PARTE III**, nel **Capitolo 5** sono stati presentati alcuni passi per lo sviluppo di uno Schema di metadati di qualità a supporto descrittivo dei contenuti di ricerca negli Archivi Istituzionali. Il medesimo schema ha offerto uno spazio per il raggruppamento e per la codifica di diversi elementi di metadati, alla luce delle:

1) "*Linee guida per la creazione e la gestione di metadati nei Repository Istituzionali*" (CRUI, 2012) che partono dal presupposto della fondamentale importanza rappresentata dai metadati negli Archivi istituzionali, che sono tra i principali strumenti che consentono di realizzare *l'accesso globale* e immediato in rete ai risultati della ricerca scientifica prodotta in una Università o in un Centro di ricerca.

Le Linee CRUI suggeriscono agli Atenei uno Schema di metadati di base (descrittivi e strutturali) e di metadati legati al contesto di appartenenza (es. informazioni amministrativo-gestionali relative ad afferenza dipartimentale tramite la Classificazione dei Settori Scientifico Disciplinari - SSD), il più possibile interoperabile nell'ambito della gestione degli Archivi (*Repository*) Istituzionali. Inoltre le Linee CRUI affrontano le problematiche legate al controllo dei Nomi, delle versioni (*versioning*) e degli identificatori univoci per ciascuna entità coinvolta nel processo di archiviazione dei dati;

2) strategie di codifica dei metadati "LODE-BD" (*Linked Open Data – enabled Bibliographic Data*) (AIMS, 2012), per rispondere al principio in virtù del quale i metadati di una piattaforma digitale devono assicurare la *migliore espressività* dei contenuti secondo le esigenze locali della descrizione e quelle informative dell'utenza, nonché la *massima visibilità, connessione, disseminazione* e il *recupero* entro e fuori i confini della piattaforma di riferimento. Le strategie di codifica "LODE-BD" - dove la "E" assume il significato di qualcosa di "*abilitato*" - affrontano particolarmente le questioni legate al controllo dei Nomi di Autori, Enti, Luoghi e alla qualificazione del metadato "Soggetto", *abilitando* i valori delle loro proprietà come *Linked Open Data* tramite l'impiego dei valori "non letterali" (URIs) dai Vocabolari controllati e Schemi di Classificazione identificabili e recuperabili sul Web.

La fusione delle indicazioni principali fornite dalle Linee Guida (CRUI) e Raccomandazioni "LODE-BD" insieme a quelle per la creazione dei metadati di qualità descritte in altri documenti ha dato luogo alla formulazione di uno nuovo Schema di metadati che potrebbe essere suggerito agli Atenei e Centri di ricerca per descrivere il loro materiale di ricerca depositato negli Archivi Istituzionali, con l'obiettivo di:

- a) misurare gli effetti di nuove strade per l'espressività e la connessione dei contenuti del materiale depositato
 - a.1) sia per mezzo delle mappature tra diversi elementi di metadati che
 - a.2) tramite l'impiego dei valori "non letterali" (URIs) di Liste di termini controllati e semantici;
- b) potenziare l'espressività delle relazioni tra i contenuti (entro e fuori l'Archivio) correlati per il contesto e la semantica e definite nel gruppo dei metadati "Relations";
- c) affiancare le risorse d'Archivio con l'uso delle licenze corrette (*Open* e *Not Open*) i cui identificatori Web possono non solo definire le modalità d'uso delle medesime risorse ma anche abilitare i valori descrittivi le proprietà di metadati "RightsHolder", "Licence," "AccessRights", "Rights" come *Linked (Open) Data*.

Il materiale di ricerca identificato tramite le Licenze *Open* rappresenta un sicuro candidato per il posizionamento dei propri contenuti in uno spazio Web – "*Linked Open Data-enabled*" chiamato *Open Knowledge*, il raggiungimento di cui è il tema di ricerca più rilevante del movimento *Open Access*, per una sempre più ampia apertura scientifica che va in parallelo con il collocamento degli elementi informativi nei territori dell'*Open Data*, dell'*Open Content*, e dell'*Open Information*, come secondo la definizione dell'*Open Knowledge Foundation* [V. "Sitografia"], un'organizzazione con lo scopo di promuovere la Conoscenza Aperta, nel tentativo di attuare una più ampia ed organizzata la condivisione di dati di diversa natura, per renderli utilizzabili, riutilizzabili, ridistribuiti per scopi sociali benefici;

- d) migliorare l'interoperabilità dei contenuti degli Archivi Istituzionali nell'ambito della loro gestione semantica.

"*It's not sufficient to apply a License in order to make a resource "Open", but the resource must also be "Open" in terms of real and effective interoperability, avoiding the use of Licenses only as something fashionable*" [De Robbio, 2011 a, online, non impaginato].

Per quanto riguarda l'allestimento della descrizione degli elementi di metadati dello Schema presentato nel Capitolo, esso ha prefissato due principali obiettivi, quello di:

- 1) dimostrare come i metadati proposti dalle "*Linee guida per la creazione e la gestione di metadati nei Repository Istituzionali*" (CRUI) possono essere affiancati dalle strategie di codifica "LODE-BD";
- 2) sviscerare alcune questioni entro una realtà che appare a tratti ricca quanto caotica, per la compilazione e qualificazione dei valori di metadati entro uno Schema tramite gli Standard di codifica della sintassi e i vocabolari controllati, laddove possibile, al fine di contribuire qualitativamente ad una rappresentazione uniforme, all'interoperabilità e allo scambio delle informazioni.

Le tematiche elencate sotto possono essere considerate tra le principali strategie a breve termine per contrastare problemi a lungo termine, per quanto riguarda l'organizzazione e la gestione dei contenuti negli Archivi Istituzionali. Esse sono:

- 1) L'elaborazione e l'implementazione degli Schemi di Metadati (dati strutturati che descrivono altri dati):

- 1.1) *ricchi*, ovvero accresciuti tramite le *mappature* tra diversi elementi di metadati e tra essi e elementi di altri schemi (es. elementi di Ontologie Web, Vocabolari controllati, Schemi semantici) che possano abilitare i metadati di base a diventare le entità *Linked Data* affinché essi possano migrare da database locali a uno spazio globale dei dati interconnessi di tipo *Linked Data* per:
 - 1.1.1) abbattere la barriera tecnologica che impedisce la libera condivisione dei dati;

- 1.1.2) elaborare meccanismi condivisi per l'*interrogazione incrociata* e l'*accesso multifocale* ai dati in maniera tale da permettere il collegamento e l'interrogazione di più dataset (contenuti) correlati per mezzo di linguaggi e protocolli standardizzati e noti;
- 1.2) aperti al *ri-utilizzo* (l'idea di un modello di metadati universalmente applicabile e valido in ogni circostanza) da diversi agenti sul Web per rintracciare, incrociare e integrare con maggiore facilità contenuti digitali di provenienza diversa, rendendoli così significativi entro un ampio contesto di condivisione sul Web Semantico. L'ambizione più alta di questo Web è la creazione di uno spazio unico di condivisione dati dove gli utenti e i Software possano trovare libertà di movimento alla ricerca di informazioni per la condivisione e il riutilizzo privo di barriere.
- 2) L'importanza di trovare soluzioni efficaci inerenti alla *garanzia dell'accesso permanente* ai materiali digitali di ricerca (incluso i dati di ricerca).
- 3) La necessità di sviluppo di un *Data Management Plan (and Process) Plans (DMPs)* per la programmazione di qualità dell'intero ciclo di vita dei contenuti digitali e dei processi in esso coinvolti.

Le tre aree tematiche appena elencate possono considerate come tre assi fondamentali da trattare e approfondire in ricerche future, per poter giungere allo sviluppo di funzionalità in grado di migliorare l'evoluzione della comunicazione dell'informazione scientifica tramite canali di disseminazione quali gli Archivi Istituzionali Open Access.

“In un mondo, come quello digitale, ad altissima velocità di mutamento, ogni previsione rischia di essere già superata nel momento stesso in cui viene espressa” [Salarelli, 2006, p.88].

Ma il compito degli archeologi del terzo millennio è RENDERE POSSIBILI FRUIZIONI DIVERSE, riaprire un SISTEMA di POSSIBILITÀ, e per farlo la prima fase è la decostruzione: azzerare ... no, certo non la costruzione, ma ciò su cui essa si basa e che incarna: UN INSIEME DI IPOTESI, DALLO SCARSO FONDAMENTO, di collocazioni e di datazioni intrise di assunzioni non esplicitate e spesso impressionistiche.

[Gambarara, Marchese, 2013, Introduzione].

Conclusions of the Thesis.

Communication of scientific information: new horizons

"Information is a difference that makes a difference"[Bateson, 1989]

This Thesis has entered in *medias res* of modalities for promotion of a new way of “socializing” research materials through the international Open Access movement, via Internet and by means of two main strategies:

- 1) Open Access Journals (“Gold Road” strategy);
- 2) Open Access Repositories /Archives (“Green Road” strategy).

These strategies or disseminating channels are considered to be complementary as they are curating aspects of publication, deposit, storage, access and consultation of research materials. Moreover these strategies fulfil fully the need of:

- a) strengthening dissemination and communication of research materials financed wholly or partially by public funds that should be of public domain, accessible to everyone, exactly accessed, consultable, used and shared via Open Access policies. This objective can be achieved by explicit placing of research material - together with its descriptive data (bibliographic data, metadata) - in the Public Domain under, for example, PDDL (Public Domain Data License) and CC0 (Creative Commons Zero) Licenses;
- b) creation of Open digital ecosystem able to embrace the communication of research information produced by different research and user communities.

PART I - Chapter 1 reconstructed the state-of-the-art regarding origins, main characteristics and statements, as well as Policies and initiatives of the Open Access movement at the international level.

PART I - Chapter 2 focused its attention on the second Open Access strategy called “Green Road” referring to the domain of Open Access Repositories (Archives) implemented on Software platforms complying with Open Archive Initiative (OAI) standard requirements. By default, these platforms are supporting Open Archive Initiative Protocol for Metadata Harvesting (OAI-PMH), thus providing interoperability (interaction) among metadata of Data and Service Providers.

The survey of the research field concerning Open Access Repositories was restricted to the observation of Institutional Repositories (IRs). These are digital platforms with a clear institutional identity (referred to Universities and Research Centers) for deposit (self-archiving), access, consultation and dissemination of research material produced by members of Universities or Research Centers.

A clear institutional identity of IRs (e.g. Institutional Repository of the University of Vienna) is a characteristic rendering this type of Repositories different from Disciplinary (Subject-based) Open Access Repositories.

After description of objectives and main features of IRs, the work forwarded to observe the field of tools and procedures that can be involved in issues concerning programming, ethics and quality of IRs. In particular, IRs were observed from the angles of:

- 1) some Guidelines and Recommendations supporting their programming, implementation and management;
- 2) their possible connections with Current Research Information Systems (CRISs). In this view, there was focused a potential task of IRs, in regard to their integration with services developed in Italy such as "Surplus Open Archive" and "U-GOV Ricerca" for integrated governance of research contents. Through these services IRs can actively participate in the assessing process of research products at the institutional (local) and national levels;
- 3) ethics related to compliance with publishing contracts and intellectual property during self-archiving (deposit) procedures and consultation of research products;

- 4) importance of the definition of Policies concerning preservation of digital content in a long term perspective, according to the conceptual model OAIS (Open Archival Information System) - ISO14721:2012 Standard;
- 5) necessity to draft (being based on IR Policies) a Data Management and Process Plan, which should contain a description of all the needed steps, processes and responsibilities involved in the qualitative management of the entire life-cycle of digital contents in a medium - long perspective. Data Management and Process Plan is considered to be a very important tool for programming of all objectives and risks involved in the managements of reserach data.

Due attention paid to the above listed aspects would surely qualitatively contribute to the quality of formation, acquisition, storage, access, dissemination and interection of reserach information. The quality of communication of IRs Content Objects (resources, their constituent parts and collections) depends on:

- a) well defined Policies;
- b) properly implemented metadata complying with widely accepted and shared Standards;
- c) good structured indexes responsible for efficient identification and classification of thema/s (Subject/s) of Content Objects qualified with values of Controlled Vocabularies and Semantic Schemes (e.g. Terminologies, Thesaurus, Classifications).

It is a good practice to normalize and harmonize "Subject" entry points of materials deposited in an IR with semantic indexes of other IRs as well as with bibliographies and catalogs of Digital Libraries and other types of authoritative digital platforms, thereby giving rise to creation and proliferation of "multi-focal" accesses to digital contents across the Web;

- d) persistent identifiers and long-term preservation strategies and metadata (e.g. OAIS: Open Archival Information System; PREMIS: Preservation Metadata Implementation Strategies), as well as well setted access rules (see Trusted Digital Repositories (TDR)). Preservation actions should guarantee immutability of characteristics of digital contents such as their reliability, authenticity and integrity for all the period in which they are considered to be useful.

While the OAIS Standard indicates with the term "Archive" a type of organization (systems and people) whose purpose is to provide and sustain information resources (packets of information objects and metadata on their support) so that they are accessible and usable in a long-term perspective by a particular community, TDR refer to the domain "Archive" through a compound term "Digital Repository". Institutional Repository is referring to both concepts: "Digital Repository" and "Archive". This means it should possess characteristics both of a digital platform for deposit of research information resources and of an Archive integrating all necessary metodologies and tools for long-term preservation of these resources. In this view, it is appropriate to refer to the Disciplinary Open Access Repository – "ArXiv" [see Website] considered to be a Best Practice for integrating functions of both "Archive" and "Digital Repository", as long as its deposited contents are accessible online already for decades.

Meanwhile, the places such as Digital Repositories and Archives are not to be confused with:

- a) Digital Libraries collecting digital information and providing access to it *but not necessarily committed to provide its long-term preservation*;
- b) Archives of Data, aimed at providing a long-term storage, but whose contents are merely statistical data sets.

The implementation of Institutional Repositories (IRs) according to the requirements of Trusted Digital Repositories (TDRs) along with well-documented Policies, procedures, practices and plans concerning digital preservation – are necessary requirements for certification (internal and external assessment) of Archive's quality. The perspective of certification must incentivize:

- 1) wide dissemination and adoption of Best Practices, Recommendations and Guidelines (see "Guidelines for Institutional Repositories" – section "Preservation", CRUI, 2009) contributing necessary instructions to effective long-term preservation programming of digital contents;

- 2) creation of new research fields investigating digital preservation requirements for IRs in the role of TDRs;
- 3) research communities, standardization bodies, ICT (Information Communication Technology) companies included Open Archives Software developers to deepen cooperatively metodological and technical issues related to digital preservation.

PART II - Chapter 3 dedicated its space to the description of a survey undertaken on a sample of infrastructures (national and international) supporting digital storage and dissemination of communication of knowledge and Free Access to research contents in internet.

In particular, the survey covered 23 infrastructures supporting Open Archive Initiative Protocol for Metadata Harvesting (OAI-PMH). There were investigated:

- a) Institutional and Disciplinary Open Access Archives accomodating digital materials having research and historic - cultural value. Most of these Archives are evaluated by online service – “Ranking Web of Repositories” [see Website, published since 2008] estimating their level of placement "among peers”;
- b) "The European Library" Portal aggregating resources by 47 European Libraries.

The objective of the sample survey was to identify:

- 1) technical requirements of the investigated platforms (Software; Statements on metadata; Availability of semantic accesses; Statistics of use of deposited materials; Declared mechanisms to support persistent access; Search/browse functionalities; programming of access levels);
- 2) additional (customed) features to expose, search/browse contents;
- 3) mechanisms connecting investigated platforms with other digital search infrastructures included channels of various Social Networks.

It is found that most of the functionalities (e.g. deposit/self-archiving, representation, search, browsing) of those infrastructures implemented on the same Software platforms disclose similar structuring and exposure. Nevertheless, a number of these infrastructures represent also customized functionalities, thus amplifying their own performance, particularly those regarding browsing, searching, presentation (display), connections of digital contents.

Good examples of customized functionalities present those platforms, which have implemented Authority Files and Semantic Schemes supporting qualification of values of “Responsible Body”, “Subject” and “Place” access points.

Qualifying metadata properties - where it is possible and/or necessary - with “non-literal” (<http://URI>) values of Controlled Vocabularies and Semantic Schemes (e.g. “Virtual International Authority File”, “Dewey Info services”) is providing an innovative approach for:

1. establishment of comparison and/or matching/mapping/cross-walking between key-words (indexes) extracted - on the basis of their occurrence - from Content Objects (information resources) withing and outside IRs;
2. creation of “multi-focal” controlled and semantic accesses to related contents withing and outside IRs.

It is clear that indexing of IRs’ information resources by controlled and semantic vocabularies and schemes presents one of the most important aspects enhancing organization of and access to information resources. On the other hand, such an approach of indexing could also present the most complex aspect of structuring information resources.

Besides being more time-consuming for indexing, the mentioned complexity can occur when, for example, in the same or in different (but technically interoperable) digital information infrastructures are implemented different KOS (Knowledge Organization Systems) referring to Authority Files, Thesaurus and Classifications. Even these tools may point to the same semantic area, they could cause strong disagreements (divergences) at linguistic, structural and semantic levels as, for example, it can occur among values with the same semantic profile but originated from disciplinary and from general Classifications (e.g. Dewey Decimal Classification, Universal Decimal Classification, Classification or the Library of Congress). One of the effective

ways to overcome possible semantic divergences is to use Controlled Vocabularies and Semantic Schemes available online as Linked Open Data (LOD) , also because many of them have been already mapped between each other, exploiting the potential of Linked Data connections based on the web addresses/identifiers <http://URIs> expressed within Resource Description Framework (RDF) (<Subject>-<Property>-<Object>) data model. Controlled Vocabularies and Semantic Schemes available online as LOD can be retrieved from: CKAN Data Hub, LOD Cloud, Library Linked Data Report (Datasets, Value Vocabularies, and Metadata Element Sets), Linked Open Vocabularies (LOV) online service.

In the conducted survey it was stressed that in IRs it would be of great advantage to base search and browsing functionalities on the values of locally developed specialized Thesaurus, whose conceptual units should be compatible or appropriated correlated with the conceptual entities of widely accepted and shared (at national and/or international levels) Thesaurus retrievable on the web as LOD. For example, at the national (Italian) level, such a specialized Thesaurus for IRs can be created in accordance to Thesaurus of New Soggettario (of the National Library of Florence) available as LOD. Such an approach will certainly:

- a) strengthen expressivity of descriptive indexes of IR not only at local level, but also within wide Linked Open Data community;
- b) favour a creation of interoperable controlled and semantic enter points in different IRs sharing the same ‘comunicating language’;
- c) amplify the network of connections among different information resources published on the web and drawn up alongside with values of Linked Open Data vocabularies and schemes. This is a key for promotion of a variety of forms of cooperation and sharing of meaning between information resources hold in digital platforms across the web.

Among the necessary factors catalyzing the value and success of IRs have been identified:

- 1) the need to follow a national/european/international best practices and guidelines to compile Policies and Plans, that can harmonize implementation and management of digital Archives infrastructures. For example, at national (Italian) level it would be a good practice to follow Guidelines CRUI, OpenAIRE, DRIVER;
- 2) the necessity to use suitable standard metadata schemas and application profiles as well as widely recognized shared metadata mappings. This task is quite significant for a better communication of digital contents within and outside IRs as well as to make possible their interaction with CRIS (Current Research Information Systems);
- 3) the integration of IRs in the administrative matters related to Research assessment at institutional and national levels;
- 4) the importance of qualifying access entry points referring to “Responsible Body” and “Subject” metadata properties with values of Thesaurus or Classifications available as Linked Open Data. Among good examples in this direction there are: VOA3R Open Access Repository, The European Library Portal, Oregon State University ScholarsArchive@OSU. A good practice enabling metadata properties to become Linked Data entities is is represented by encoding strategies of the Linked Open Data Enabled – Bibliographical Data (LODE-BD) Recommendations (VOA3R Open Access Repository, AGRIS platform);
- 5) the implementation of unique Persistent Identifiers and formats supporting long-term preservation and access of digital contents;
- 6) the implementation of value-added services such as "Statistics of use of deposited materials", "Critical collaborative comments", "Share", “Like” and other functionalities typical to Social Networks, “Forum” and “Blog” virtual spaces);
- 7) the importance of development of links among content objects inside and outside of IRs ‘confines’. This approach can surely provide a high promising way for development of a common search virtual space "Global Scientific Data Space."

PART II - Chapter 4 presented a ToolKit based on the state of the art of some aspects in support of the technical profile of Digital Archives. In particular, some matters and reported

references to relevant theoretical observations, projects, tools, methodologies and practical solutions have been briefly introduced in conjunction with the following sections:

- 1) Software platforms;
- 2) Metadata in support of digital contents;
- 3) Authority control of metadata values;
- 4) Association of metadata with semantic values;
- 5) Storage and long-term reservation;
- 6) Digitization;
- 7) Resource discovery: paradigms of searching and browsing.

Contents developed for these sections have not a claim to be exhaustive. They can be rather points from which to develop and compile further strategies supporting technical and non technical aspects of IRs.

PART III - Chapter 5 investigated some steps and frameworks for the development of qualitative metadata elements that can support description of digital contents. Hereby, a metadata schema describing IRs' research materials is developed. It is created in accordance with:

- a) "Guidelines for metadata creation and management in Institutional Repositories" (CRUI, 2012) assuming the fundamental importance of metadata elements for description, access and retrieval of IRs' content objects. The Guidelines CRUI suggests a Scheme based ultimately on Dublin Core metadata (descriptive and structural), and on metadata related to administrative University context (e.g. administrative information related to Departmental affiliation through Classification of Scientific-Disciplinary Sectors – SSD). According to Guidelines CRUI, metadata schema elements for IRs must be as much interoperable as possible in the context of IRs and outside their borders, focusing on CRIS (Current Research Information Systems). In addition, the Guidelines CRUI address the issues related to the authority control of Names (regarding "Responsible Body" metadata), of versions and of unique identifiers for each entity involved in the process of storing/archiving data;
- b) encoding metadata strategies of "Linked Open Data - enabled Bibliographic Data (LODE-BD) Recommendations" (AIMS, 2012), in order to respond to the principle whereby the metadata of any digital platform must ensure the best expressivity of its contents according to the local descriptive and user needs. Metadata should guarantee maximum visibility, access, dissemination, retrieval and connection of digital contents within and outside platforms in which they reside.

The encoding strategies "LODE-BD" - where the "E" assumes the meaning of something "Enabled" – are addressing particularly issues related to the control and semantization of metadata values referred to "Responsible Body", "Place", "Subject", "Relation" metadata elements. The main objective of "LODE-BD" is to enable bibliographical metadata to become Linked (Open) Data entities, making charge of "non-literal" (<http://URIs>) values from Controlled vocabularies and Classification schemes identifiable and retrievable on the Web as Linked Open Data as well as of mark-ups from "Schema.org" Web Ontology, and of elements from "BIBO" Ontology expressing bibliographical data on the Semantic Web.

The fusion of the main recommendations provided by the "Guidelines CRUI" and "LODE-BD" documents, together with suggestions of other authority documents in support of metadata creation guided towards the formulation of a new metadata Schema. This one can be proposed for use in IRs of Universities and Research Centers, in order to describe research materials.

The objective of the presented schema is to:

- a) measure some new effects of expression and connections of the deposited IRs research materials through:
 - a.1) mappings (cross-walking) among different metadata elements;

- a.2) use of “non-literal values” expressed by unique web addresses (<<http://URIs>>) corresponding to “literal values” of Controlled Vocabularies and Semantic schemes;
- b) enhance the expressiveness of relations (see “Relation” metadata group) between digital content units within and outside IRs;
- c) evaluate the choice of Licenses (Open and Not Open) to assign a correct use of deposited resources and to enable metadata properties such as "RightsHolder", "Licence", “AccessRights”, " Rights” to acquire characteristics of Linked (Open) Data entities.

The research material published under Open Licenses and according to Linked Data publishing strategies is an enabled candidate to enter into the information space of connected data called “Linked Open Data Cloud”, as well as to efficiently interact with spaces such as “Open Knowledge”, “Open Data”, “Open Content”, “Open Information”, according to Open Knowledge Foundation’s [see Website] definitions. The achievement of this objective should be primary for the research material promoted through Open Access movement, in order to contribute - in open and shared way - to the creation of cross-domain research Open Knowledge, in an attempt to implement increasingly broad and organized sharing of various kinds of Open research datasets (contents), to make them interoperable, reusable, and efficiently redistributed among different communities of users.

“It is not sufficient to apply a License in order to make a resource "Open", but the resource must also be "Open" in terms of real and effective interoperability, avoiding the use of Licenses only as something fashionable” [De Robbio, 2011 a, online].

The description of metadata elements of the presented metadata Schema has two main objectives:

- a) to show how metadata proposed by the Guidelines CRUI can be drawn up alongside with encoding "LODE-BD" strategies;
- b) to go into depth on some issues regarding compilation and qualification of metadata values by means of encoding standard syntax and controlled vocabularies (where necessary and possible), in order to improve the quality of uniform representation, interoperability and exchange of information resources.

The topics listed below can be considered as main short-term strategies to tackle possible long-term issues in Open Access IRs:

1) The necessity to develop and implement Metadata Schemas (considered as structured data describing other data):

- 1.1) richened, or enhanced through mappings (cross-walks) among different metadata elements and among them and elements of other schemes (e.g. Web Ontologies, Controlled vocabularies, Semantic schemes), with "non-literal values" based on web addresses (<<http://URIs>>). In this view metadata would be able to migrate easily from local databases to a global space of interconnected datasets - Linked Open Data;
- 1.2) able to tear down the technological barriers preventing free sharing of data;
- 1.3) enabled to support mechanisms for “cross-querying” and “multi-focal access” to data allowing the aggregation of multiple datasets (content), jointly by means of standardized and widely shared web protocols and data frameworks;
- 1.4) which can be opened to re-use (the idea of an universally applicable and valid in all circumstances metadata model) by different agents. Metadata schemas should be easily traced, intersected and integrated, making significant their elements within a broader context of data on the Semantic Web. The highest ambition of this web is the creation of a single data sharing space where agents (users and Software) can be free to search, share and reuse digital contents without any legal or technological barriers.

2) The importance to find effective solutions to guarantee permanent access to digital research materials (including research data).

The continuing investment into in-house preservation systems is contributing to the lack of interoperability and fragmentation of resources into “digital silos”. Not transferrable and interchangeable stand-alone solutions lead to fragmentation and do not offer economies of scale. Shared solutions for creation, storage and use of digital resources, including the e-Infrastructures, will become the major component of the future Knowledge Economy.

3) The need to develop a Data Management and Process Plans (DMPs) to qualitatively program the entire life-cycle of digital contents and the processes on their support.

The three thematic areas listed above could be considered as three fundamental axis that can be addressed to future research. It is important to find further new solutions for improvement of communication of scientific information through dissemination channels such as Open Access Institutional Repositories, accelerating and facilitating research content sharing, reusability, exchange and providing qualitative permanent access to research materials across different communities operating in and contributing to the highly-interactive ecosystem of Linked Open Data.

BIBLIOGRAFIA

Adorni, Giovanni, Coccoli, Mauro, Vercelli, Gianni, Vivinet, Giuliano. Topic Maps e XTM per l'e-learning, in "Journal of e-Learning and Knowledge Society", Vol. 3, n. 3, settembre 2007, <http://www.academia.edu/3216606/Topic_Maps_e_XTM_per_le-learning>

Aguillo, Isidro. F., Ortega, José. L., Fernández, Mario, Utrilla, Ana.M. Indicators for a webometric Ranking of Open Access Repositories, in "Scientometrics", 82 (3): 477-486, 2009, <<http://digital.csic.es/bitstream/10261/32190/1/Ranking%20of%20Repositories.pdf>>

Aliprandi, Simone. Apriti Standard. Interoperabilità e formati aperti per l'innovazione tecnologica, Ledizioni 2010, ISBN: 9788895994345: <http://www.ledizioni.it/stag/wp-content/uploads/woocommerce_uploads/2011/11/9788895994345_content.pdf>

----- Capire il Copyright, Percorso guidato nel diritto d'autore, Ledizioni 2012, Collana: CopyLeft Italia, ISBN: 9788867050116: <http://www.ledizioni.it/stag/wp-content/uploads/woocommerce_uploads/2012/06/Aliprandi_CapireCopyright_webDEF.pdf>

Aliprandi, Simone, Creative Commons: a user guide, Ledizioni, 2011, ISBN: 9788895994550: <http://www.aliprandi.org/cc-user-guide/aliprandi_cc_user_2.pdf>

Allison, Julie, Johnston, Pete, Powell, Andy: A Dublin Core Application Profile for Scholarly Works, in "Ariadne", 30 January 2007, <<http://www.ariadne.ac.uk/issue50/allinson-et-al/>>

Awre, Chris and Swan, Alma, Linking Repositories: scoping the development of cross-institutional user-oriented services. *OCLC Systems & Services: International Digital Library Perspectives*, 2012 (last modified), <http://eprints.soton.ac.uk/263962/>

Altman Micah. Metadata and Metrics to support Open Access Monographs. Workshop on Open Access and Scholarly Books, [Berkman Center](#) and [Knowledge Unlatched](#), 2013: <<http://drmltman.wordpress.com/2013/06/20/metdata-and-metrics-to-support-open-access-monographs/>>

Baek, Jae-Eun, Sugimoto, Shigeo. Facet Analysis of Archival Metadata Standards to Support Appropriate Selection, Combination and Use of *Metadata Schemas*, in DCMI global meetings & Conferences, DC-2011, the Hague: <<http://dcevents.dublincore.org/IntConf/dc-2011/paper/view/39>>

Baethge, Christopher, Impact Factor—a Useful Tool, but Not for All Purposes, in "Dtsch Arztebl Int.", 2012 April; 109(15): 267–269: <<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3345343/#R3>>

Baker, Tomas. La traduzione dei dati nel linguaggio del Web Semantico, in "JLIS.it", 2012 a, <<http://leo.cilea.it/index.php/jlis/article/view/6308/7866>>

----- "Linking your resources to Data Web", Webinar, 4th December, 2012 b, <<http://aims.fao.org/community/metadata/blogs/did-you-miss-first-linked-open-data-aims-webinar-tom-baker-now-you-can-acce>>

----- Bechhofer, Sean, Isaac, Antoine, Miles, Alistair, Schreiber, Guus, Summers, Ed, *Key choices in the design of Simple Knowledge Organization System (SKOS)*, in “*Web Semantics: Science, Services and Agents on the World Wide Web*”, 10 May 2013, <<http://www.sciencedirect.com/science/article/pii/S1570826813000176>>

Ball, Alexander, *Preservation and Curation in Institutional Repositories*, DCC, UKOLN at University of Bath, 2010: <<http://www.dcc.ac.uk/resources/briefing-papers/technology-watch-papers/curation-repositories>>

Barton, Jane, Currier, Sarah, Hey Jessie M. N. *Building Quality Assurance into Metadata Creation: an Analysis based on the Learning Objects and e-Prints Communities of Practice*, in “*Cataloging & Classification Quarterly*”, 47 (2009), issue 3-4, <<http://eprints.erpanet.org/83/01/paper60.pdf>>

Bateson, Gregory, Bateson, Mary Catherine. *Dove gli angeli esitano. Verso un'epistemologia del sacro*, Milano, Adelphi, Milano, 1989.

Beall, Jeffrey, *Predatory publishing is just one of the consequences of gold Open Access*, in “*Learned Publishing*”, Volume: 26 Issue: 2, April, 2013: <<http://www.researcherid.com/ProfileView.action?SID=V2x4lloHYOTtR5Vazob&returnCode=ROUTER.Success&queryString=KG0UuZjN5WmoB8bI5snMN0aQCwppw5hg8xtymdrD9m60%253D&SrcApp=CR&Init=Yes>>

Bergamin, Giovanni, Lucarelli, Anna. *The NUOVO SOGGETTARIO as a service for the Linked Data world*, in “*JLIS.it*”, Vol.4, n.1 (Gennaio/January 2013), <<http://leo.cilea.it/index.php/jlis/article/view/5474>>

Berners-Lee, Tim, *LinkedData. DesignIssues*. Status: personal view only. Editing status: imperfect but published, 2006, <<http://www.w3.org/DesignIssues/LinkedData.html>>

Bizer, Christian, “*Expert Report on Linking Data & Publications*”, in *GRDI2010: A vision for global Research data infrastructures*, 2013, <<http://151.1.219.218/b43d3f37-bd5d-4144-9779-b27a0ca3d1d5.pdf>>

Broughton, Vanda. *Una classificazione per il 21' secolo. Principî e struttura della Classificazione bibliografica BLISS*. University college London. School of Library, archive and information studies. – Trad. it. di Gnoli, Claudio, 2004, <<http://www.aib.it/aib/contr/broughton1.htm>>

Caracciolo, Caterina, Stellato, Armando, Morshed, Ahsan, Johannsen, Gudrun, Rajbahndari, Sachit, Jaques, Yves, Keizer, Johannes. *The AGROVOC Linked Dataset*, in “*Semantic Web – Interoperability, Usability, Applicability* “ (an *IOS Press Journal*), 05/18/2012, <<http://www.semantic-web-journal.net/content/agrovoc-linked-dataset>>

Caruso, Francesco. *Sperimentazione e analisi di tecniche di Riconoscimento dei Caratteri (OCR) per la digitalizzazione di documenti. Relazione di Stage*. Università di Roma Sapienza, 2007, <https://www.cnipa.gov.it/site/files/Refo2_RS_11.0_Caruso.pdf>

Caplan, Priscilla. *Capire PREMIS*. The Library of Congress. Trad. it. Di Iorio, Angela, Fondazione Rinascimento Digitale, 2009, <http://www.loc.gov/standards/premis/Understanding-PREMIS_italian.pdf>

Cassella Maria, La gestione dei diritti nei progetti di digitalizzazione: il pubblico dominio e le opere orfane, in “JLIS.it”, Vol. 4, n. 2 (Luglio/July 2013), <<http://www.google.it/url?sa=t&rct=j&q=&esrc=s&frm=1&source=web&cd=1&cad=rja&ved=0CDEQFjAA&url=http%3A%2F%2Fleo.cilea.it%2Findex.php%2Fjlis%2Farticle%2Fdownload%2F8797%2F8061&ei=uPYUvH3OdLT4QTR5IDgDw&usg=AFQjCNGIf1CxzFtdEhrHG-IUBo-NYkbEpg&sig2=DfrTPRtCcF-4vdRF010h-Q>>

----- Bozzarelli, Oriana “Open Access, Archivi Istituzionali e valutazione della ricerca”, Politecnico di Torino, la settimana Open Access, 24 ottobre 2011, <http://eprints.rclis.org/bitstream/10760/16775/1/OA2011_AIBnotizie.pdf>

Caso, Roberto, *Scientific knowledge unchained: verso una Policy dell’università italiana sull’Open Access*, in “Trento Law and Technology Research Group”, Maggio 2013, <<http://eprints.biblio.unitn.it/archive/4155/>>

----- Galimberti Paola, La Legge italiana sull’Accesso Aperto agli articoli scientifici: l’inizio di un percorso normativo, in “ROARS Returns on Academic Research”, 6 novembre 2013: <<http://www.roars.it/online/la-legge-italiana-sullaccesso-aperto-agli-articoli-scientifici-linizio-di-un-percorso-normativo/>>

----- Il “Signore degli anelli” nel ciberspazio: controllo delle informazioni e *Digital Rights Management*”, in Borghi, M. Montagnani, M. L. (curr.), “*Proprietà digitale*” – *Diritto d’autore, nuove tecnologie, e Digital Rights Management* – Atti del Convegno, l’Università Commerciale Luigi Bocconi, Egea, Milano, 18 novembre 2005, Versione 1.0 - settembre 2006, <http://eprints.biblio.unitn.it/archive/00001133/01/Roberto_Caso.DRM.Signore_degli_anelli.pdf>

Castelli, Donatella, Manghi, Paolo, Thanos, Costantino. Comunicazione scientifica: la nuova frontiera, in “DigItalia”, Anno VII, Numero 2 -2012, ISSN 1972-6201: <<http://www.google.it/url?sa=t&rct=j&q=openaire%20enhanced%20publication%20%20C3%A8%20una%20pubblicazione%20&source=web&cd=4&ved=0CEEQFjAD&url=http%3A%2F%2Fdigitalia.sbn.it%2Farticle%2Fdownload%2F576%2F408&ei=umuFUZfZDYWg4gSXgIHIDg&usg=AFQjCNFvNvg37h21QbHNCCtFgh21-933Rg>>

Castellucci, Paola, Tempo e massa: una nuova energia nella comunicazione scientifica, in “Bollettino AIB”, Vol 51, N° 3 (2011), <<http://bollettino.aib.it/article/view/5428>>

Chan, Lois Mai, Zeng, Marcia Lei, *Metadata Interoperability and Standardization: A Study of Methodology*. Part I. Achieving Interoperability at the Schema Level, in “D-Lib Magazine”, 12 (2006) 6, <<http://www.dlib.org/dlib/june06/chan/06chan.html>>

Clarke, Olivia. *The Learning Federation digital content and the pre-service education sector*, in “Curriculum Leadership: An electronic Journal for leaders in education“, Volume 6 , Issue 12, 2008, <http://www.curriculum.edu.au/leader/the_learning_federation_digital_content.22998.html?issueID=11387>

Corlosquet, Stéphane, Clark, Lin, *The Semantic Web, Linked Data and Drupal*, Part 1: Expose your data using RDF, in “Developer Works”, 12 Apr 2011, <<https://www.ibm.com/developerworks/web/library/wa-rdf/>>

Cullen, Rowena, Chawner, Brenda., *Institutional Repositories, Open Access, and Scholarly Communication: A Study of Conflicting Paradigms* , in “JAL” (“The Journal of Academic

<<https://www.zotero.org/marlened/items/itemKey/TJFEPMTU>; <http://www.slis.tsukuba.ac.jp/aliep2009/proceedings/Papers/a37.pdf>>

Cyganiak, Richard, Stenzhorn, Holger, Delbru Renauld, Decker, Stefan, Tummarello, Giovanni, *Semantic Sitemaps: Efficient and Flexible Access to Datasets on the Semantic Web*, in: *Proceedings of the 5th European Semantic Web Conference, ESWC, 2008*, <<http://richard.cyganiak.de/2008/papers/sitemaps-eswc2008.pdf>>

DC (Dublin Core) – 2013 - “Linking to the future”. Conference Proceedings - Papers, Project Reports and Posters (Introduction to Linked Open Data; Introduction to Ontology Concepts and Terminology; Metadata Provenance; Data enrichment and transformation in the LOD context: poor and popular vs. rich and lonely – can we achieve both?; Why Schema.org?; Provenance and Annotations for Linked Data..): <<http://dcevents.dublincore.org/index.php/IntConf/dc-2013/schedConf/presentations>>; <<http://dcpapers.dublincore.org/pubs/issue/view/165>>

DC (Dublin Core) – 2011 – “Metadata Harmonization: Bridging Languages of Description”. Conference Proceedings: <<http://dcevents.dublincore.org/index.php/IntConf/dc-2011/schedConf/presentations>>

De Castro, Paola, La POLICY per l'Open Access dell'Istituto Superiore di Sanità: strategie e innovazione. Convegno “*Open Access, digital preservation e deposito legale: Policy, progetti e servizi per la ricerca*”, 8 maggio 2008, <<http://bice.cnr.it/images/pdf/De%20Castro%20CNR%208%20maggio%202008.pdf>>

Debating Open Access, ed. by Nigel Vincent and Chris Wickham, The British Academy, 2013: <<http://www.britac.ac.uk/openaccess/debatingopenaccess.cfm>>

De Robbio, Antonella. Un formato europeo per i dati di biblioteche, musei e archivi, in “*il Bo: il giornale dell'Università degli studi di Padova*”, 3 settembre 2013: <<http://www.unipd.it/ilbo/content/un-formato-europeo-i-dati-di-biblioteche-musei-e-archivi>>

----- *Gold vs Green: le differenti vie dell'Accesso Aperto*, in “*il Bo: il giornale dell'Università degli studi di Padova*”, 1 ottobre 2013: <<http://www.unipd.it/ilbo/content/gold-vs-green-le-differenti-vie-dell%E2%80%99accesso-aperto>>

----- La ricerca condivisa: nasce *Zenodo*, il *Data Repository* europeo, in “*Giornale dell'Università degli studi di Padova*”, 24 maggio 2013, <<http://www.unipd.it/ilbo/content/la-ricerca-condivisa-nasce-zenodo-il-data-repository-europeo>>

----- *Open Access* come ‘quinta libertà’ nei progetti di ricerca e innovazione, in “*il Bo: il giornale dell'Università degli studi di Padova*”, 28 marzo 2013: <<http://eprints.rclis.org/19570/1/5.libertaOA.pdf>>

----- Editoria scientifica 2013: nuove tendenze oltre la “primavera accademica”, in “il Bo: il giornale dell’Università degli studi di Padova”, 6 marzo 2013: <<http://www.unipd.it/ilbo/content/editoria-scientifica-2013-nuove-tendenze-oltre-la-%E2%80%9Cprimavera-accademica%E2%80%9D>>

----- Chi sorveglia gli editori predatori, in “IL BO: il giornale dell’Università degli studi di Padova”, 20 febbraio 2013: <<http://www.unipd.it/ilbo/content/chi-sorveglia-gli-editori-predatori>>

-----Biblioteche, quanto si spende davvero per le riviste scientifiche?, in “IL BO: il giornale dell’Università degli studi di Padova”, 12 luglio 2013, <<http://www.unipd.it/ilbo/content/biblioteche-quanto-si-spende-davvero-le-riviste>>

-----Auto-archiviazione per la ricerca: problemi aperti e sviluppi futuri, in “Bibliotime”, anno IV, n.3 (novembre 2003), <<http://didattica.spbo.unibo.it/bibliotim/num-vi-3/derobbio.htm>>

-----Archivi aperti e comunicazione scientifica, Università degli Studi di Napoli Federico II, ClioPress, 2007, <<http://www.fedoa.unina.it/1093/1/derobbio.pdf>>

De Robbio, Antonella, Katzmayer, Michael, Gestione e organizzazione di un Archivio Aperto internazionale Disciplinare: il caso E-LIS, in “AIDAInformazioni: rivista di Scienze dell’informazione”, 2009, vol. Anno 2, <<http://eprints.rclis.org/13201/>>

De Robbio, Antonella, Giacomazzi, Silvia *Dati aperti con LODE*, in “Bibliotime”, anno XIV, numero 2 (luglio, 2011 a): <<http://didattica.spbo.unibo.it/bibliotime/num-xiv-2/derobbio.htm#nota91>>

De Robbio, Antonella. OL4OD: Licenze Aperte per dati aperti, in “JLIS.it.”, Vol. 2, n. 2 (Dicembre/December, 2011 b): <<http://eprints.rclis.org/16443/1/4695-5561-2-PB.pdf>>

----- Forme e gradi di apertura dei dati. I nuovi alfabeti dell’Open Biblio tra scienza e società, in *Biblioteche oggi*, 30 (2012), n. 6, <<http://www.bibliotecheoggi.it/content/201200601101.pdf>>

----- Metadati: parola chiave per l’accesso alla biblioteca ibrida. Convegno “La biblioteca ibrida: Verso un servizio informativo integrato”, Milano, 14-15 marzo 2002, <<http://eprints.rclis.org/bitstream/10760/4001/3/stelline%26metadati.pdf>>

De Paoli, Nilde, Mornati, Susanne. SURplus Open Archive (OA): il Repository Istituzionale a servizio della ricerca, in “Bollettino del Cilea”, n. 113 (giugno 2009), p. 16-19, <<http://www.google.it/url?sa=t&rct=j&q=&esrc=s&frm=1&source=web&cd=1&cad=rja&ved=0CDMQFjAA&url=http%3A%2F%2Fbollettino.cilea.it%2Farticle%2Fdownload%2F6935%2F6530&ei=NvkUUoDyIMSD4gTP2oGgAQ&usq=AFQjCNEIkyt4UxJWqRGv464T7gY7vTAYuQ&sig2=vvf-VK2Cr9jprOBzYN7H1Q>>

Delfanti, Alessandro. *The Politics of Open Science*, Pluto Press, 7 May, 2013, <<http://www.youtube.com/watch?v=szbE1OynKxU>>

Di Donato, Francesca, Müller, Susanne. Biblioteche Digitali semantiche. Il progetto Burckhardtsource.org, «Bibliotime», anno XVI, numero 1 (marzo 2013), <<http://eprints.rclis.org/19038/1/didonato.htm>>

Direttiva 2013/37/UE del Parlamento Europeo e del Consiglio del 26 giugno 2013 che modifica la direttiva 2003/98/CE relativa al riutilizzo dell'informazione del settore pubblico, Gazzetta ufficiale dell'Unione europea, 27.6.2013:

<<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2013:175:0001:0008:IT:PDF>>

Disegno di legge 26 settembre 2013. Approvato dal Senato della Repubblica. Conversione in legge, con modificazioni, del decreto-legge 8 agosto 2013, n. 91, recante disposizioni urgenti per la tutela, la valorizzazione e il rilancio dei beni e delle attività culturali e del turismo: <http://download.repubblica.it/pdf/2013/cultura/disegno_legge_tutela.pdf>

Doerr, Martin, Gradmann, Stefan, Henniecke, Steffen, Isaac, Antonie, Meghini Carlo, Van de Sompel, Herbert, *Europeana Data Model (EDM)*, in *IFLA 2010 Libraries and the Semantic Web. Session 149, Information Technology, Cataloguing, Classification and Indexing with Knowledge Management*, World Library and Information Congress, 2010, Gothenburg, Sweden, <<http://conference.ifla.org/past/ifla76/149-doerr-en.pdf>>

Dulong de Rosnay, Melanie, De Martin, Juan Carlos "*The Digital Public Domain: foundations for an Open Culture*", ed. by, Lightning Source for Open Book Publishers, 2006, <http://www.communia-association.org/wp-content/uploads/the_digital_public_domain.pdf>

Dunsire, Gordon, *Linked Data, Libraries and the Semantic Web*, in "*Library science talk*", 12-13 March 2012, <http://www.nb.admin.ch/aktuelles/ausstellungen_und_veranstaltungen/00726/01611/03953/03958/index.html?lang=en>

Dunsire, Gordon, Willer, Mirna. *Initiatives to make Standard library metadata models and structures available to the Semantic Web*, Conferenza IFLA 149. *Information Technology, Cataloguing, Classification and Indexing with Knowledge Management*, 10-15 August 2010, Gothenburg, Sweden, <<http://conference.ifla.org/past/ifla76/149-dunsire-en.pdf>>

Duranti, Luciana *Records in the Cloud: Towards InterPARES Trust*, Fondazione Rinascimento Digitale, 2012, <http://www.rinascimento-digitale.it/conference2012/paper_ic_2012/duranti_paper.pdf>

Eke, Helen Nneka . *Digitizing resources for University of Nigeria repository: Process and challenges*, in "Webology", Volume 8, Number 1, June, 2011, <<http://www.webology.org/2011/v8n1/a85.html>>

Faiella, Feliciana, *Digital preservation: Le problematiche della conservazione nel contesto dell'informazione digitale*, Biblioteche oggi - Ottobre 2000, <<http://www.bibliotecheoggi.it/2000/20000806401.pdf>>

Fay, Ed, *Repository Software Comparison [DSpace 1.6, EPrints 3.2.0 and Fedora 3.3]: Building Digital Library Infrastructure at LSE*, in "Ariadne", Issue 64, July 2010, <<http://www.ariadne.ac.uk/issue64/fay/>>

Faiella, Feliciana. *DIGITAL PRESERVATION: Le problematiche della conservazione nel contesto dell'informazione digitale*, in Biblioteche oggi, Ottobre 2000, p.64,

<http://www.bibliotecheoggi.it/2000/20000806401.pdf>

Ferwerda, Eelco, Snijder, Ronald, Adema, Janneke. Report: A project exploring Open Access Monograph publishing in the Netherlands”, 10 november, 2013:

<http://www.surf.nl/en/publicaties/Pages/ReportAprojectexploringOpenAccessmonographpublishingintheNetherlands.aspx>

Galimberti, Paola , Archivi istituzionali e valutazione della ricerca: l'importanza dei dati e della loro validazione, in “Bollettino del CILEA”, N. 116 dicembre 2010 a:

<http://bollettino.cilea.it/article/view/7002>

Galimberti, Paola, Towards a new scenary for research assessment: the Institutional Repository (AIR) of Milan University, in “JLIS.it”, 2010 b: <http://eprints.rclis.org/handle/10760/15194>

Galimberti, Paola, *Verso un nuovo scenario per la valutazione della ricerca: potenzialità dell'Open Access e limiti imposti dal contesto*, in “JLIS.it”, Vol. 1, n. 1 (Giugno/June 2010), p. 87–110, <http://leo.cilea.it/index/jlis/article/viewFile/16/23>

Gambarara, Daniele . Thesauri, mappe semantiche, ontologie. Problemi semantici e costruzioni concettuali, a c. di Guarasci R., *Dal documento all'informazione*, Milano, ITER, 2008, pp. 249-276.

----- Marchese Maria Pia, Ed., Guida per un'Edizione Digitale dei manoscritti di Ferdinand De Saussure, a c. di, Collana: Studi e ricerche – 117, Edizioni dell'Orso, Srl A Alessandria, ISBN: 978-88-6274-478-2, 2013: <http://www.ediorso.it/guida-per-un-edizione-digitale-dei-manoscritti-di-ferdinand-de-saussure.html>

Gasevic, Dragan, Djuric, Dragan, Devedzic, Vladan, Damjanovi, Violeta. *Converting UML to OWL Ontologies*, Proceedings of the 13th international World Wide Web conference, New York, 2004, <http://portal.acm.org/citation.cfm?id=1013539>

Getaneh, Alemu, Brett, Stevens, Penny, Ross, *Semantic Metadata Interoperability in Digital Libraries: A Constructivist Grounded Theory Approach*, in ACM/IEEE Joint Conference on Digital Libraries, Ottawa (Canada), 13 June 2011:

<http://eprints.rclis.org/15829/>

Giglia, Elena. *Academic Social Networks: It's Time to Change the Way We Do Research*, in “European journal of physical and rehabilitation medicine”, 47.2 (2011): 345–350. Print

Ghirardi Nicola, Linked Data in Digital Library, CINECA, Bologna, 23 maggio 2013:

<http://www.slideshare.net/ghirardinicola/presentazioneeventodllinked-data-indigitallibrary11>

Glaser, Hugh, Halpin Harry. The Linked Data strategy for global identity, in “IEEE Internet Computing”, 16.2 (2012): 68–71. (Cit. on p. 9): <http://eprints.soton.ac.uk/333924/>

Gnolli, Claudio “Organizzazione della conoscenza”: *Web-publish or perish*, in “AIDAinformazioni“, Anno 28 gennaio-giugno, Num. 1-2/2010, presentazione: <http://www.dimat.unipv.it/gnoli/works.php?cl=7h>

Goddard, Lisa, Byrne, Gillian. *Linked Data tools: Semantic Web for the masses*, “First Monday”, 15 (2010), n. 11:

<http://www.uic.edu/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/3120/2633>

Gnoli, Claudio, Scognamiglio, Carlo *Ontologia e organizzazione della conoscenza*, Lecce: Pensa Multimedia, 2008.

Gordon Dunsire, *Interoperability and semantics in RDF representations of FRBR, FRAD and FRSAD*, 2010, <<http://www.gordondunsire.com/pubs/pres/InteropSemanticsRDFFRBR.ppt>>

Gorman, Michael, *La biblioteca come valore. Tecnologia, tradizione e innovazione nell'evoluzione di un servizio*. Trad. di Matteo Barucci, a c. di Mauro Guerrini, postf. Alberto Petrucciani. Udine, Forum, 2004.

Greenberg, Jane. *Understanding Metadata and Metadata Schemes*. In "Cataloging & Classification Quarterly", 40 (2005) 3/4:
<<http://www.ils.unc.edu/mrc/pdf/greenberg05understanding.pdf>>

----- *Metadata and the World Wide Web*, in "Encyclopedia of Library and Information Science", New York: Marcel Dekker, Inc., 2003,
<<http://www.ils.unc.edu/mrc/pdf/greenberg03metadata.pdf>>

----- Davenport W. Robertson, Semantic Web construction: an inquiry of authors' views on collaborative metadata generation, in "Proceedings of the International Conference on Dublin Core and Metadata for e-Communities", 2002, 45-52,
<<http://www.bncf.net/dc2002/program/ft/paper5.pdf>>

Guarasci, Roberto, Le Viste Documentali, in *Conservare il Digitale*, a c. di Pigliapoco S. *Archivistica Informatica*, EUM, 2010, pp. 177-191.

----- Che cos'è la documentazione, a c. di Guarasci R., *Dal documento all'informazione*, Milano, ITER, 2008.

Guerrini, Mauro. Introduzione al Seminario "Global interoperability and Linked Data in libraries", in "JLIS.it", Vol 4, N.1, 2013, *special issue: "Global Interoperability and Linked Data in Libraries" (the papers delivered at the Seminar Global Interoperability and Linked Data in Libraries, held on 18th and 19th June, 2012 at the University of Florence)*,
<<http://leo.cilea.it/index.php/jlis/issue/view/536>>

----- Crupi, Gianfranco. L'Open Access: work in progress, in "JLIS.it": *Special issue: Open Access milestones*, Vol.3, n.2 (Dicembre/December 2012), pp.1-4:
<<http://www.ledizioni.it/prodotto/jlis-vol-3-n-2-2012-special-issue-open-access-milestones/>>

----- Possemato Tiziana, *Linked data: un nuovo alfabeto del Web Semantico*, in "Biblioteche oggi", aprile 2012, <<http://www.bibliotecheoggi.it/content/201200300701.pdf>>

----- Possemato Tiziana, Cosa sono i *Linked Data*, in "JLIS.it", Vol. 4, n. 1 (Gennaio/January 2013), pp.9-10, <<http://leo.cilea.it/index.php/jlis/article/view/6305/7892>>

----- Gli Archivi Istituzionali. *Open Access*, valutazione della ricerca e diritto d'autore, Editrice Bibliografica, Milano, 2010 a, <<http://hdl.handle.net/10760/15609>>

----- *Open Access: il ruolo delle biblioteche*. Il Gruppo di lavoro CRUI sui metadati con la collaborazione di Marchitelli Andrea (CILEA), seminario “*Open Access e conoscenza aperta: quali vantaggi per chi fa ricerca?*” 2010 b, <http://dspace-unito.cilea.it/bitstream/2318/710/1/Guerrini_Open_Access_il_ruolo_delle_biblioteche.pdf>

----- Ventura, Roberto, *Problemi dell’editoria universitaria oggi: il ruolo delle University Press e il movimento a favore dell’Open Access*, in “*Dalla pecia all’e-book. Libri per l’università: stampa, editoria, circolazione e lettura*”. Atti del convegno internazionale di studi, Bologna, 21-25 ottobre 2008 (a c. di G.P. Brizzi, M. G. Tavoni), Bologna: CLUEB, 2009, p. 665-670, <<http://eprints.unifi.it/archive/00002004/01/50Guerrini.pdf>>

----- Ventura, Roberto, *Biblioteca e biblioteconomia, ovvero, Del rapporto fra bibliotecario e docente di biblioteconomia*, a c. di: Innocenti, P., Cavallaro, C. *Una mente colorata : studi in onore di Attilio M. Caproni per i suoi 65 anni*, Vecchiarelli, Roma, 2007, <<http://eprints.unifi.it/archive/00001763/>>

----- Tillett, Barbara, Sardo, Lucia. *Authority control. Definizioni ed esperienze internazionali*. Atti del convegno internazionale, Firenze, 10-12 febbraio 2003. Firenze: Firenze University Press, Associazione Italiana Biblioteche, 2003: <<http://www.fupress.com/Archivio/pdf/4383.pdf>>

----- *Open Access: il ruolo delle biblioteche*. Il Gruppo di lavoro CRUI sui metadati con la collaborazione di Marchitelli Andrea (CILEA), seminario “*Open Access e conoscenza aperta: quali vantaggi per chi fa ricerca?*” 2010, <http://dspace-unito.cilea.it/bitstream/2318/710/1/Guerrini_Open_Access_il_ruolo_delle_biblioteche.pdf>

Guédon, Jean-Claude. *The “green” and the “gold” road to Open Access: the case for mixing and matching*, in “*Serials Review*”, 30 (2004) 4, p.315-328

----- *Mixing and matching the Green and the Gold road - Take two*, in “*Serials review*”, March 2008, 34 (1), p. 41-51, <<http://eprints.rclis.org/handle/10760/11791>>

Gunning, Tashina. *Metadata creation at Institutional Repositories*, in “*PNLA Quarterly*”, 75:4, 2011, <<http://unllib.unl.edu/LPP/PNLA%20Quarterly/gunning75-4.pdf>>

Hall, Steven, *Widening access to research information: collaborative efforts towards transitions in scholarly communications*, presentato in *Berlin 7 Open Access Conference: Open access reaching diverse communities*, 2-4 December 2009, La Sorbonne, Paris, France, <<http://www.berlin7.org/IMG/pdf/hall.pdf>>

Harnad, Stevan, *Scholarly Skywriting and the Prepublication Continuum of Scientific Inquiry, reprinted*, in “*Current Contents*” 45: 9-13, November 11 1991, <<http://cogprints.org/1581/>>

----- Les, Carr, Swan, Alma, Sale, Arthur and Bosc, Helene, *Open Access repositories: maximizing and measuring research impact through university and research-funder Open Access self-archiving mandates*. *Wissenschaftsmanagement* 4, Juli/August, 36-41, 2009, <<http://eprints.ecs.soton.ac.uk/16616/>>

Harpring, Patricia, Baca, Murtha, *Introduction to Controlled Vocabularies: Terminology for Art, Architecture, and Other Cultural Works*, Series Editor © 2010 J. Paul Getty Trust,

<http://www.getty.edu/research/publications/electronic_publications/intro_controlled_vocab/index.html>

Hogenaar Arjan. Enhancing Scientific Communication through Aggregated Publications Environments. *Ariadne* Issue 61, 30-October, 2009, <<http://www.ariadne.ac.uk/issue61/hogenaar/>>

Hatop Götz , Integrating Linked Data into Discovery, in *Code4Lib Journal*, Issue 21, <<http://journal.code4lib.org/articles/8526>>

Heath, Tom, Bizer, Christian. *Linked Data: evolving the Web into a global data space*, in “Synthesis lectures on the semantic web: theory and technology”, Morgan & Claypool, 2011, <<http://linkeddatabook.com/editions/1.0/>>

Höckner, Markus, Budroni, Paolo PHAIDRA - A Repository-Project of the University of Vienna, in “INFOtheca“, № 1, vol XII, August 2011, <http://infoteka.bg.ac.rs/PDF/Eng/2011-1/INFOTHECA_XII_1_August2011_23a-32a.pdf>

Hogenaar, Arjan, *Enhancing Scientific Communication through Aggregated Publications*, in “*Ariadne*”, October 2009, Issue 61, <<http://www.ariadne.ac.uk/issue61/hogenaar/>>

Iacono, Antonella. Verso un nuovo modello di OPAC. Dal recupero dell’informazione alla creazione di conoscenza, in *JLIS.it*. Vol. 4, n. 2 (Luglio/July 2013), <<http://leo.cilea.it/index.php/jlis/article/download/8903/8069>>

IFLA Working Group on Functional Requirements and Numbering of Authority Records (FRANAR), Requisiti funzionali per i dati di autorità. Un modello concettuale, a c.d. Glenn E. Patton, Edizione italiana. A cura dell’Istituto Centrale per il Catalogo Unico delle biblioteche italiane e per le informazioni bibliografiche (ICCU), Roma, 2010, <http://www.ifla.org/files/assets/cataloguing/frad/frad_2009-it.pdf>

IFLA Working Group on the Functional Requirements for Subject Authority Records (FRSAR) . Functional Requirements for Subject Authority Data (FRSAD). *A Conceptual Model*, ed. Marcia Lei Zeng, Maja Žumer, Athena Salaba, 2010, <<http://www.ifla.org/files/assets/classification-and-indexing/functional-requirements-for-subject-authority-data/frsad-final-report.pdf>>

Innocenti, Perla, Ciuccarelli, Paolo. Sistemi conoscitivi per il design. Una proposta metodologica. Il caso DesignNet, POLI.design , Milano, 2004

Jackson, Joab, *Tim Berners-Lee: Machine-readable Web still a ways off*, in “CGN”, October 2009, <<http://gcn.com/articles/2009/10/30/berners-lee-semantic-web.aspx>>

Jarar, Mustafa, Dikaiakos, Marios. *A Quesry Formulation Language for the Data Web*, in “IEEE Transaction on Knowledge and Data Engineering”, Vol. 24(5): 783-798 (2012): <<http://www.informatik.uni-trier.de/~ley/pers/hd/j/Jarrar:Mustafa>>

Jeffery, Keith, Corson-Rikert, Jon. *euroCRIS and VIVO. Part II Cooperation as Strategic Partners*, 2012: <http://www.vivoweb.org/files/presentations/12Fri/euroCRIS_LOD_and_%20VIVO.pdf>

Johnson, Thomas, Boock, Michael. *Linked Data Services for Theses and Dissertations*, Proceedings of the 15th International Symposium on Electronic Theses and Dissertations, Lima, Peru, 12-14 September 2012, <http://ir.library.oregonstate.edu/xmlui/bitstream/handle/1957/32977/JohnsonBoock2012_ETD.pdf?sequence=1>

Kaplan ,Andreas, Haenlein Michael, *Users of the world, unite! The challenges and opportunities of social media*, in “Business Horizons”, Vol. 53, Issue 1, 2010, p. 59-68

Keßler, Mirjam. *Linked Open Data of the German National Library*, in: *ECO4r Workshop LOD of DNB. Deutsche National Bibliothek*, 2010, <http://www.eco4r.org/workshop2010/eco4r_workshop2010_mirjam_kessler.pdf>

Keith, Jeffery, Asserson, Anne. *Institutional Repositories and Current Research Information Systems*, in “New Review of Information Networking”, 14(2), 2008: <<http://www.tandfonline.com/doi/abs/10.1080/13614570903359357>>

Koutsomitropoulos, Dimitrios A., Alexopoulos, Andreas D., Solomou, Georgia D., Papatheodorou, Theodore S., *The Use of Metadata for Educational Resources in Digital Repositories: Practices and Perspectives*, in “D-Lib Magazine”, January/February 2010, Vol.16, Num. ½, <<http://www.dlib.org/dlib/january10/kout/01kout.html>>

Lagoze, Carl, *The Warwick Framework: A container architecture for aggregating sets of metadata*, in “D-Lib Magazine”, July/August, 1996, <<http://www.dlib.org/dlib/july96/lagoze/07lagoze.html>>

Lavoie, Brian, Gartner, Richard. 'Preservation Metadata', DPC Technology Watch Report, DPC Charles Beagrie Ltd, 13-03 May 2013, <<http://dx.doi.org/10.7207/twr13-03>>

Le Hénaff, Diane, Sicilia Miguel-Angel, Gavrilut, Laura. *Open Repositories as Social Networks: the case of VOA3R 2012*, in “Open Repositories International” Conference (Edinburgh, United Kingdom), 2012 <<http://prodinra.inra.fr/?locale=fr#!ConsultNotice:163304>>

Lucarelli, Anna, Franci, Luciana, Motta, Marta, Rolle, Massimo. *The Nuovo Soggettario Thesaurus: structural features and web application projects*, in IFLA 2009 Satellite meetings in Florence, Florence 17-21 August, <<http://www.bncf.firenze.sbn.it/pagina.php?id=209&rigamenu=Italian%20National%20Bibliography%20-%20News>>

Lucarelli. Anna, *Semantic Authority Control and the New Soggettario.*, 2003 . In *International Conference Authority Control: Definition and International Experiences*, Florence, February 10-12- 2003, <<http://eprints.rclis.org/4167/>>

Lueg, Christopher, Bidwell, Nicola. *Berrypicking in the Real World: A Wayfinding Perspective on Information Behavior Research*, in Proceedings of the American Society for Information Science and Technology, Volume 42, Issue 1, 2005, p.1., <<http://onlinelibrary.wiley.com/doi/10.1002/meet.14504201241/full>>

Manghi, Paolo, Bolikowski, Lukasz, Manola, Natalia, Schirrwagen, Jochen, Smith, Tim. *OpenAIREplus: the European Scholarly Communication data Infrastructure*. In “D-Lib magazine”, September/October 2012, Vol. 18, Num. 9/10,

<http://www.dlib.org/dlib/september12/manghi/09manghi.html>>

Marcato, Andrea. *La Ricerca a Faccette: Aspetti teorici e pratici*. Tesi. Università degli Studi di Padova, 2010, <http://tesi.cab.unipd.it/26030/1/La_Ricerca_a_Faccette.pdf>

Marchitelli, Andrea. *Verso la circolazione dei saperi pubblici. DSpace e gli Archivi Istituzionali*, Roma, 2007, <http://eprints.rclis.org/bitstream/10760/10872/1/presentazione_marchitelli.pdf>

Maurits van der Graaf, Kwame van Eijndhoven, *The European Repository Landscape: Inventory Study into the Present Type and Level of OAI-Compliant Digital Repository Activities in the EU*, Amsterdam University Press, 2008, <<http://dare.uva.nl/document/93725>>

Meschini, Federico. *eContent: tradizionale, semantico o 2.0?*, in “Delivering e-content: contenuti elettronici alla portata degli utenti”, Milano, 2006, p. 72-86, Roma, AIDA Edizioni, <<http://dspace.unitus.it/handle/2067/162>>

Michetti, Giovanni. *Una norma per la gestione documentale: il modello ISO 15489*, in *DigItalia*, Vol 1., 2007, <<http://digitalia.sbn.it/article/view/412>>

Mietchen Daniel, Maloney Chris, and Moskopp Nils Dagsson, “*Inconsistent XML as a Barrier to Reuse of Open Access Content*”, in Conference (JATS-Con) Proceedings 2013: <<http://www.ncbi.nlm.nih.gov/books/NBK159964/>>

Moffat, Malcolm, *Marketing' with Metadata: How Metadata Can Increase Exposure and Visibility of Online Content with "Best Practice" Metadata*, 2006, in *New Review of Information Networking*, vol.12, issue 1-2, 2006, <<http://www.tandfonline.com/doi/abs/10.1080/13614570601136248>>

Mornati, Susanna. *L'Accesso Aperto negli Atenei Italiani: raccomandazioni*, in “*Raccomandazioni per lo sviluppo dell'editoria elettronica negli atenei italiani*”, a c.d. Patrizia Cotoneschi, Giancarlo Pepeu, CRUI, 2006, <<http://digital.casalini.it/8884534577>>

Mornati, Susanna. *CRIS e repository istituzionali delle pubblicazioni: la proposta SURplus di CILEA*. In “*BOLLETTINO DEL CILEA*”, N. 116, *dICEMBRE 2010*: <http://www.google.it/url?sa=t&rct=j&q=&esrc=s&frm=1&source=web&cd=5&cad=rja&ved=0CFIQFjAE&url=http%3A%2F%2Fwww.researchgate.net%2Fpublication%2F49601759_CRIS_e_repository_istituzionali_delle_pubblicazioni_la_proposta_SURplus_di_CILEA%2Ffile%2F9fcd50addee62f68a.pdf&ei=GVwOUsWpKcOg4gTvo4DQAg&usq=AFQjCNH4xAqSymtzVaKEWz93pVL2_kdyhQ&sig2=2kMJWGCY00LIE7APXcv9-w>

Mornati, Susanna. *SURplus, innovazione ed efficacia nei CRIS*, *Current Research Information Systems*, in “*BOLLETTINO DEL CILEA*”, N. 117, *giugno 2011*, <<http://www.google.it/url?sa=t&rct=j&q=&esrc=s&frm=1&source=web&cd=1&cad=rja&ved=0CDEQFjAA&url=http%3A%2F%2Fbollettino.cilea.it%2Farticle%2Fdownload%2F6973%2F6568&ei=GVwOUsWpKcOg4gTvo4DQAg&usq=AFQjCNFnp9tt5QnIMsBbphGzLXErhIqReg&sig2=wQySnrwT41phXSHdlpe7FQ>>

Nicholas, David, Rowlands, Ian, Watkinson, Anthony, Brown, David, Hamid R. Jamali. *Digital repositories ten-years on: what do scientific researchers think of them and how do they use them?*, in “*Learned Publishing*”, 25, 2012, pp. 195–206 <http://ciber-research.eu/download/20120620-Digital_repositories_ten_years_on.pdf>

Nigel, Vincent , Wickham, Chris (ed). *Debating Open Access (a collection of 8 articles by leading academics and publishers reflecting on the challenges of Open Access publication for humanities and social sciences)*, British Academy, 2013,
<<http://www.britac.ac.uk/openaccess/debatingopenaccess.cfm>>

Nilsson, Mikael. *From Interoperability to Harmonization in Metadata Standardization: Designing an Evolvable Framework for Metadata Harmonization*. PhD Dissertation. KTH School of Computer Science and Communication. Stockholm, Sweden, 2010,
<<http://kmr.nada.kth.se/papers/SemanticWeb/FromInteropToHarm-MikaelsThesis.pdf>>

O'Brien, Margaret, Using the OBOE Ontology to Describe Dataset Attributes, in “LTER Databits“, 11/14/2010, <<http://databits.lternet.edu/fall-2010/using-oboe-ontology-describe-dataset-attributes>>

Oh, Sam Gyun. «MARC, FRBR and RDA: The Topic Maps Perspective», Slides, 2008,
<<http://www.topicmaps.com/tm2008/oh.pdf>>

O'Reilly, Tim. *What Is Web 2.0*. Online su *O'Reilly Network*, 30 settembre 2005,
<<http://oreilly.com/web2/archive/what-is-web-20.html>>

Paci, Giulio, Pedrazzi, Giorgio, Turra, Roberta, *Wikipedia based semantic metadata annotation of audio transcripts*, in “Proceedings of the 11th Workshop on Image Analysis for Multimedia Interactive Services” WIAMIS 2010, IEEE Computer Society, Los Alamitos, CA, USA,
<<http://ieeexplore.ieee.org/xpl/articleDetails.jsp?reload=true&arnumber=5617667>>

Pasqui, Valdo “Accesso persistente per oggetti digitali ad Accesso Libero”, *Briefing Paper* per “Digital Preservation Europe”, *University of Florence*, 2012:
<http://www.digitalpreservationeurope.eu/publications/briefs/accesso_persistente.pdf>

Pattueli M. Cristina, Miller Matt, Lange Leanora, Fitzell Sean, Li-Madeo Carolyn *Crafting Linked Open Data for Cultural Heritage: Mapping and Curation Tools for the Linked Jazz Project*, in “Code4Lib” Journal, Issue 21, <<http://journal.code4lib.org/articles/8670>>

Pepper, Steve. *Topic Maps and the Semantic Web*, in online post “*Topic map and all that*”, 2009,
<http://topicmaps.wordpress.com/2008/05/11/topic-maps-and-the-semantic-web/>

Pievatolo, Maria Chiara, “Dalla Libertà della penna alla Libertà della Tastiera: Accesso Aperto, diritto d'autore e mediazione editoriale: una lettura illuministica”. Video - Lezione, Sistema Bibliotecario del Politecnico di Torino, 17 maggio 2013:
<<http://rivoluzionedigitale.polito.it/2013-pievatolo>>

Poynder, Richard, *Ten years after*, in “Information today”, vol. 21, n. 9 (October 2004), p. 1-2,
<<http://www.infotoday.com/IT/oct04/poynder.shtml>>

Ricolfi. Marco. Consume and Share: Making Copyright Fit for the *Digital Agenda*, in “The Digital Public Domain: Foundations for an Open Culture”, ed. by M.Dulong De Rosnay, J.C. De Martin, OpenBook Publishers, 2012, pp.47-92, <http://www.communia-association.org/wp-content/uploads/the_digital_public_domain.pdf>

Ridi, Riccardo, *La biblioteca come ipertesto: verso l'integrazione dei servizi e dei documenti*. Milano: Editrice bibliografica, 2007. 352 p.

Rovella, Anna. *Produzione ordinamento e gestione dei documenti*, a c. di R. Guarasci, *Dal documento all'informazione*, Milano, ITER, 2008, pp. 161-183.

----- Guarasci Roberto, Murzaku Alexander, Cavarretta Giuseppe Alfredo, *La classificazione automatica dei documenti*, in "Archivi", 2009, Vol. IV, n. 1, pp. 75-82.

----- Marre' Giovanni, *Strutturazione dell'informazione e integrazione della conoscenza*, in "AIDA Informazioni", 2009, Vol. 1-2 2008, pp. 253-262.

Taverniti Maria, Guarasci Roberto, Rovella Anna, *Archivi Digitali e profili formativi nell'UE*, in "Books seem to me pestilent things", Cavallaro C. (a c. di), Manziana (Roma): Vecchiarelli, 2011, Vol. 4, pp. 1295-1302.

Rudasill, Lynne M., Dorta-Duque, Maria Elena, *Open Access and Digital Libraries: Social Science Libraries in Action / Acceso Abierto y Bibliotecas Digitales : Bibliotecas de Ciencias Sociales en Acción*, Berlin/Munich: De Gruyter Saur, ISBN 978-3-11-028102-6, (IFLA Publications; Nr 158), 2013:

<<http://www.ifla.org/publications/ifla-publications-series-158>>

Rusu, Delia, Fortuna, Blaž, Mladenia, Dunja *Automatically Annotating Text with Linked Open Data*, in: "WWW2011 workshop: Linked Data on the Web (LDOW2011)", Hyderabad (India), 2011,

<<http://events.linkedata.org/ldow2011/papers/ldow2011-paper09-rusu.pdf>>

Saffady, William. *Digital library concepts and technologies for the management of library collections: an analysis of methods and costs*. In "Library Technology Reports", 31 (1995) 3, pp. 221-380. Available for purchase online at:

<<http://cat.inist.fr/?aModele=afficheN&cpsidt=2485292>>.

Safari, Mehdi, *Metadata and the Web*, in *WebOlogy*, vol. 1, n.2, December, 2004,

<http://eprints.rclis.org/bitstream/10760/7318/1/Metadata_and_the_Web.pdf>

Salarelli Alberto, Tammaro Anna Maria, *La biblioteca digitale*, Seconda edizione, Editrice Bibliografica, Milano, 2006.

Serrai, Alfredo, *Natura, elementi e origine della bibliografia*, Roma, Bulzoni, 2009

Salo, Dorothea. *Name authority control in institutional repositories*, in "Cataloging and Classification Quarterly", 47:3/4 (April 2009):

<<http://minds.wisconsin.edu/handle/1793/31735>>

Sandhaus, Evan, Larson, Rob. *First 5000 tags released to the Linked Data Cloud*, in "New York Times", 29 ottobre 2009, <<http://open.blogs.nytimes.com/2009/10/29/first-5000-tags-released-to-the-linked-data-cloud/>>

Segnini A., Vespucci, L. *I matematici "tosati"*, in "AIDA informazioni", (ottobre-dicembre, n.4), 2004, <<http://hdl.handle.net/10760/10283>>

Sequeda, Juan. “*Consuming Linked Data*”. Proc. of Semantic Technology Conference, 2011: <http://www.slideshare.net/juansequeda/consuming-linked-data-semtech2010>

Shearer, Kathleen, Schmidt, Birgit Lossau, Norbert “*Licensing Revisited: Open Access Clauses in Practice*”, in “LIBER“ 41st Annual Conference 27-30 June 2012, Tartu, Estonia, <http://www.utlib.ee/liber2012/index.php?id=programme&item=acc_papers>

Šimek Pavel, Le Hénaff Diane, Sicilia Miguel -Angel. *Use of shared central virtual open access agriculture and aquaculture repository*. In “*Open Repositories*” 2012 Conference (Edinburgh, United Kingdom), <<http://prodinra.inra.fr/?locale=en#ConsultNotice:163509>>

Šimek, Pavel. *Using Metadata Description for Agriculture and Aquaculture Paper*. *Agris on-line Papers*, in “*Economics and Informatics*”, 4.4 (2012): <http://online.agris.cz/files/2012/agris_on-line_2012_4_simek_vanek_ocenasek_stoces_vogeltanzova.pdf>.

Solodovnik, Iryna *Open Access: un panorama internazionale*, in “*Biblioteche Oggi*”, Vol.XXX, 2012, n.6, pp.52-57, <<http://www.bibliotecheoggi.it/content/n201206.html>>

Solodovnik, Iryna *Strategie di metadati nella ricognizione di aspetti strutturali e semantici in rete*, in “*Culture del testo e del documento le discipline del libro nelle Biblioteche e negli Archivi*”, fasc. 38 (Maggio-Agosto 2012), n. 2 della nuova serie, <<http://www.unitus.it/discutedo/culture.htm>>

----- Comunicazione e ricerca semantica di contenuti informativi tra Metadati, *Linked open Data* e Ontologie, in “*Bibliotime*“, XIV, num. 2 (luglio 2011), <<http://hdl.handle.net/10760/15966>>

----- *Bibliographic Data towards the Semantic Web: a review of key issues and recent experiences*, in “*Information World Journal*“, 2012, pp.17-56, <http://www.bd.org.tr/index.php/bd/article/view/3>;

----- Uno sguardo sul futuro semantico dell’universo bibliografico, in “*AIB: Associazione Italiana Biblioteche*“, Vol. 52, n.3 (2012), <<http://aibstudi.aib.it/issue/current/showToc>>

----- Il Progetto VOA3R: la costruzione del sociale e semantico per la ricerca e la valutazione dei contenuti nei Repository Open Access, in “*Bibliotime*“, XVI, 1, 2013, <<http://www.aib.it/aib/sezioni/emr/bibtime/num-xvi-1/solodovnik.htm>>

Stracke, Christian M., Le Hénaff, Diane, Sicilia, Miguel-Angel. *Community-based research, knowledge sharing and publication facilitated by VOA3R*, in “*European Magazine FOCUS EUROPE*“, 2012: <http://www.c.enter-network.eu/fileadmin/CENTER_UP/Focus_Europe_IV_October_2012.pdf>

Stevenson, Jane. *Content Architecture: Exploiting and Managing Diverse Resources*, in “*Ariadne*“, issue 60, 29 July 2009, <<http://www.ariadne.ac.uk/issue60/isko-2009-rpt>>

Suber, Peter, *Open Access*, MIT Press, 2012, <http://mitpress.mit.edu/sites/default/files/titles/content/9780262517638_Open_Access_PDF_Version.pdf<https://plus.google.com/109377556796183035206/posts/QGDsbZd7DtE>>

- *Breve Introduzione all'Accesso Aperto*, traduzione italiana di Susanna Mornati.
<<http://www.aepic.it/docs/OA/brief-italian.htm>>
- Subirats, Imma; Zeng, Marcia L. *Meaningful Bibliographic Metadata (M2B)*. Rome: Food and Agriculture Organization of United Nations, 2012, <<http://aims.fao.org/fr/metadata/m2b>>
- *LODE-BD Recommendations 2.0: How to select appropriate encoding strategies for producing Linked Open Data(LOD)-enabled bibliographic data*. Rome: Food and Agriculture Organization of United Nations, 2012, <<http://aims.fao.org/lode/bd>>
- Keizer, Johannes, *Metadata Approaches for Shareable and LOD-enabled Bibliographic Data from Open Repositories*, in DCMI global meetings & Conferences, DC-2011, the Hague: <<http://dcevents.dublincore.org/intconf/dc-2011/paper/view/32>>
- Sutton, Caroline, “*Is Free Inevitable in Scholarly Communication? The Economics of Open Access*”, in “*College & Research Libraries News*”, © 2011 Caroline Sutton, <<http://crln.acrl.org/content/72/11/642.full>> (L’articolo ripercorre la storia dell’evoluzione delle riviste dalla carta al formato elettronico, e dal formato elettronico all’*Open Access*, sottolineando come il fatto che le riviste scientifiche siano in forma digitale le sottopone a principi economici diversi rispetto a quelle tradizionali a stampa).
- Swan, Alma, UNESCO Policy Guidelines on the Development and Promotion of Open Access. UNESCO, Paris, 2012: <<http://unesdoc.unesco.org/images/0021/002158/215863e.pdf>>
- Sharing knowledge: open access and preservation in Europe. Conclusions of a strategic workshop, Brussels, 25-26 November 2010, 2011, <http://ec.europa.eu/research/science-society/document_library/pdf_06/oa-preservation-2011_en.pdf>
- Swan, Alma, Institutional repositories - now and next. In: *University Libraries and Digital Learning Environments* (eds Penny Dale, Jill Beard and Matt Holland), Ashgate Publishing, 2011, <<http://eprints.ecs.soton.ac.uk/21471/>>
- Modelling scholarly communication options: costs and benefits for universities. A report for the JISC, 2010, <<http://eprints.ecs.soton.ac.uk/18584/>>
- New metrics for research outputs: overview of the main issues. Technical Report, 2010, <<http://eprints.ecs.soton.ac.uk/16986/>>
- Brown, Sheridan, To share or not to share: Publication and quality assurance of research data outputs. Report to the Research Information Network, 2008, <<http://eprints.ecs.soton.ac.uk/16742/>>
- Open Access and the progress of science. *American Scientist* 95 (3), May-June 2007, pp.197-199, <<http://eprints.ecs.soton.ac.uk/13860/>>
- Awre Cris, Linking UK repositories: Technical and organisational models to support user-oriented services across institutional and other digital repositories: Scoping study report, 2006: <http://www.jisc.ac.uk/uploaded_documents/Linking_UK_repositories_report.pdf>
Appendix:

http://www.jisc.ac.uk/uploaded_documents/Linking_UK_repositories_appendix.pdf

Tammaro, Anna Maria, Nuove prospettive per la comunicazione scientifica: l'esperienza della Firenze University Press, in "Biblioteche Oggi" (2001) 4, p.22-34.

Tarrant, David, O'Steen, Ben, Brody, Tim, Hitchcock, Steve. *Using OAI-ORE to Transform Digital Repositories into Interoperable Storage and Services Applications*, in "Code4Lib", Issue 6, 2009-03-30, <http://journal.code4lib.org/articles/1062>

Tartaglia, Stefano *Controllo d'autorità e linguaggi di indicizzazione per soggetto*. In *International Conference Authority Control: Definition and International Experiences*, Florence, February 10-12, 2003, <http://eprints.rclis.org/4190/>

Taylor, Coyle, *An introduction to Metadata, Information Access Service*. University of Queensland Library, Australia, 2003, <http://www.library.uq.edu.au/iad/ctmeta4.html>

Trigari, Marisa. *L'interoperabilità semantica nei metadati*. Seminario nazionale di studi, Roma, 2007, <http://www.otebac.it/getFile.php?id=134>

Tunkelang, Daniel, *Faceted Search*, in "Synthesis Lectures on Information Concepts, Retrieval, and Services", Google Inc., 2009, 80 pp., <http://www.morganclaypool.com/doi/abs/10.2200/S00190ED1V01Y200904ICR005>

Vanni, Luca, Fini, Antonio. *Learning object e metadati. Quando, come e perché avvalersene*, Trento, Erickson, 2004

Vassallo, Salvatore *Descrizioni archivistiche e Web Semantico: un connubio possibile?*, in "JLIS.it", Vol. 1, n. 1 (Giugno/June 2010), <http://leo.cilea.it/index/jlis/article/viewFile/27/37>

Villazón-Terrazas, Boris, Corcho, Oscar, *Methodological guidelines for publishing Linked Data*, Wood, David (Ed) *Linking Government Data*, 2011, http://delicias.dia.fi.upm.es/wiki/images/7/7a/07_MGLD.pdf

Vorndran, Angela, Botte, Alexander, "An Analysis and evaluation of existing methods and indicators for quality assessment of scientific publications". European Educational Research Quality Indicators (EERQI), 2008, http://www.eerqi.eu/sites/default/files/Analysis_and_evaluation_of_existing_methods_and_indicators.pdf

Weston, Paul Gabriel, *I metadati e il catalogo elettronico. Interoperabilità di contenuti e servizi digitali: metadati, standard e linee guida*. Seminario nazionale di studi, Roma, 2007, <http://www.otebac.it/getFile.php?id=131>

----- Galeffi, Agnese. Il Controllo d'autorità come raccordo fra sistemi descrittivi dei beni culturali: prospettive e progetti in ambito bibliotecario, in "Archivi e Computer: automazione e beni culturali", 4 (2004), n.2, p.85-116

-----Galeffi, Agnese. "Il Controllo d'autorità come raccordo fra sistemi descrittivi dei beni culturali: prospettive ed esperienze", Milano, Sala Convegni dell'Archivio di Stato, 26 novembre 2003, <http://www.archiviodistatomilano.it/uploads/autornov03/pdf/weston-galeffi-tx.pdf>

White, Wendy. “*Institutional repositories: contributing to institutional knowledge management and the global research commons*”, in *4th International Open Repositories Conference*, 2009, <<http://eprints.soton.ac.uk/48552/>>

Wilson, Neil, *Linked Data Prototyping at the British Library*, Slides, 2010, <<http://talis-linkeddata-libraries.s3.amazonaws.com/Linked%20Data%20Prototyping.pdf>>

Zareen Syed, Finin, Tim *Creating and Exploiting a Hybrid Knowledge Base for Linked Data*, in: *Agents and Artificial Intelligence*, Joaquim Filipe, Ana Fred, Bernadette Sharp (Eds.). Revised Selected Papers Series: Communications in Computer and Information Science, Springer, 129 (2011), p. 3-24, <http://ebiquity.umbc.edu/paper/html/id/535>

Zumer, Maja, Zeng, Marcia, Salaba, Athena. *FRBR: A Generalized Approach to Dublin Core Application Profiles*, in *DCMI International Conference on Dublin Core and Metadata Applications, DC-2010, Pittsburgh Proceedings*, 2010, <<http://dcpapers.dublincore.org/pubs/article/view/1024>>

SITOGRAFIA

ACCESSO APERTO. Che cos'è? F.A.Q. & Answers di Maria Chiara Pievatolo, 6 ottobre 2013:
<<http://www.roars.it/online/accesso-aperto-che-cose-f-a-q-answers/>>

ACCESSO APERTO è legge – in Germania, 2013:

<<http://minimacademica.wordpress.com/2013/09/20/laccesso-aperto-e-legge-in-germania/>>

“L'autore di un contributo scientifico che ha avuto origine nell'ambito di un'attività di ricerca e *insegnamento finanziata almeno per metà da fondi pubblici* ed è pubblicato in una collezione che esce periodicamente almeno due volte l'anno *ha il diritto* - anche se ha concesso all'editore o al curatore un diritto d'uso esclusivo – di *rendere pubblicamente accessibile*, dopo la scadenza di dodici mesi dalla prima pubblicazione, il contributo nella versione del manoscritto accettato, fin tanto che non serva a uno scopo commerciale. La fonte della prima pubblicazione deve essere indicata”.

A Bibliographic Framework for the Digital Age. Library of Congress, 2013:

<<http://www.loc.gov/marc/transition/news/framework-103111.html>>; <<http://bibframe.org/>>

Acquisition & Collection Development Section (ACD). IFLA, 2013. Trad it. Punti chiave per lo sviluppo delle collezioni come risorsa elettronica: <<http://www.ifla.org/files/assets/acquisition-collection-development/publications/electronic-resource-guide-it.pdf>>

ADDENDUM per l'autore al contratto editoriale di pubblicazione:

<http://paduaresearch.cab.unipd.it/docs/SPARC%20AUTHOR%20ADDENDUM%20traduzione.pdf>

Agenda Digitale Europea, “Gazzetta ufficiale dell'Unione europea“ del 26 agosto 2010:

<[http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:52010DC0245R\(01\):EN:NOT](http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:52010DC0245R(01):EN:NOT)>

AgriVIVO. Portale di ricerca: <<http://www.agrivivo.net>; <http://www.agrivivo.net/page/terms-use>>

AGROVOC Thesaurus come Linked Open Data - LOD (Search Terms; Browse Hierarchies; Concept Scheme; Maintenance; Copyright & Download; Linked Data; Web Services): ><http://aims.fao.org/standards/agrovoc/about>>

AIMS (Agricultural Information Management Standards): <<http://aims.fao.org/>>

Alcuni canali web per l'indicizzazione del material scientifico:

- DART-Europe E-theses Portal: <<http://www.dart-europe.eu/basic-search.php>>
- DOAJ (Directory of Open Access Journals): <<http://www.doaj.org/>>
- DOAR (Directory of Open Access Repositories): <<http://www.opendoar.org/>>
- DRIVER (Digital Repository Infrastructure Vision for European Research): <<http://www.driver-repository.eu/>>
- Europeana (una mega-biblioteca digitale europea): <<http://www.europeana.eu/>>
- OAISTER (Union Catalog of digital resources): <<http://www.oclc.org/oaister.en.html?urlm=168646>>
- JSTOR: <<http://www.jstor.org/>>

- NARCIS (National Academic Research and Collaborations Information System): <<http://www.narcis.nl/>>
- OpenAire (Open Access Infrastructure for Research in Europe): <<http://www.openaire.eu/it>>
- RePEc (Repository of Research Papers in Economics): <<http://repec.org/>>
- ROAR (Registry of Open Access Repositories): <<http://roar.eprints.org/>>
- Scholarius: <<http://www.scholarius.com/>>
- Scirus: <<http://www.scirus.com/>>
- VOA3R (Virtual Open Access Agriculture & Aquaculture Repository): <<http://voa3r.eu/>>
- WorldCat (un mega-catalogo online delle collezioni di biblioteche a livello mondiale): <<http://www.worldcat.org/>>

ANSI/NISO Z39.88 - 2004: The OpenURL Framework for Context-Sensitive Services
<http://www.niso.org/standards/standard_detail.cfm?std_id=783>

APARSEN (*Alliance Permanent Access to the Records of Science in Europe Network*) Project:
<<http://www.alliancepermanentaccess.org/index.php/aparsen/>>

Appunti per la definizione di un set di metadati gestionali-amministrativi e strutturali per le risorse digitali: <<http://www.iccu.sbn.it/metaAG1.pdf>>

AquaBrowser Library software (la ricerca è supportata da una rete di associazioni semantiche: associazioni tra termini, varianti ortografiche, termini provenienti da Thesauri e vocabolari controllati, traduzioni in altre lingue prodotte nella fase di configurazione ed indicizzazione dei dati sul database del sistema:
<<http://win.cenfor.net/products/database/bowker/AquaBrowserBrochure.pdf>>

Archeological Open Data, Opening the Past. MapPapers 3-II, 2012:
<http://mappaproject.arch.unipi.it/wp-content/uploads/2011/08/Pre_atti_online3.pdf>

Archiviazione documentale: 2C Solution:
<<http://www.archiviazione-documentale.it/acquisizione-archiviazione-documenti-cartacei-e-analogici.php>>

Archivio della didattica dell' Università degli studi di Bologna: <<http://almadl.cib.unibo.it/>>

Archivio digitale fotografico Open Access MDC, Digital Memory of Catalunya:
<<http://mdc1.cbuc.cat/index.php>>

Archivio Disciplinare arXiv.org: <<http://arxiv.org/>>

Archivio Istituzionale dell'Università di Bergamo AISBERG: <<http://aisberg.unibg.it/>>

Archivio Istituzionale della Ricerca dell'Università degli studi di Milano AIR:
<<http://air.unimi.it>>

Archivio Istituzionale della Ricerca dell'Università degli studi di Milano AIR - "Diritto d'Autore": <<http://air.unimi.it>>

Archivio Istituzionale “EconStor” dell’Università Humboldt (Berlino) con la certificazione DINI: <<http://www.econstor.eu/>; <http://edoc.hu-berlin.de/>>

Archivio istituzionale dell’Università di Trieste, OpenstarTs: <<http://www.openstarts.units.it/dspace/?locale=it>>

ArgoUML Open Source UML Modeling Tool: <<http://argouml.tigris.org/>>

Associazione Italiana Biblioteche (AIB) a favore dell’Accesso Aperto alla letteratura scientifica, 2006: <<http://www.aib.it/aib/cen/open.htm>>

AIB-DFP, “Stato e necessità della documentazione di fonte pubblica in rete”, 2007: <www.aib.it/dfp/c0711d.htm3>

AIDALampi: l’informazione in rete sul mondo della documentazione avanzata, agosto 11, 2009, <<http://www.aidalampi.it/?p=678>>

Archivio Storico della Royal Society, <<http://royalsociety.org/news/Royal-Society-journal-archive-made-permanently-free-to-access/>> (La Royal Society è l’editore scientifico più antico al mondo, avendo pubblicato la prima edizione delle Philosophical Transactions of the Royal Society nel 1665. Nell’Archivio si trovano documenti rari come il primo articolo di Isaac Newton, le opere giovanili di Charles Darwin, gli esperimenti di Benjamin Franklin; CIBER; Comitato universitario, Basi dati ed editoria in rete, 2 novembre 2011: <<http://bib04.caspur.it/cibernewsletter/?p=14588>>)

ARIEL: Portale della didattica online dell’Università di Milano: <<http://ariel.unimi.it/User/Default.aspx>>

ArXiv (Archivio disciplinare per i fisici): <<http://arxiv.org/>>

ArXiv - SWORD/APP Deposit API User's Manual: <http://arxiv.org/help/submit_sword>

Benchmark DP Project – Systematic approach to assess and compare digital preservation processes, systems, and organizational capabilities Systematic approach to assess and compare digital preservation processes, systems, and organizational capabilities: <<http://benchmark-dp.org/>>

Benefits of the Linked Data Approach: <<http://www.w3.org/2005/Incubator/ld/wiki/Benefits>>

Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities, 2003: <www.zim.mpg.de/openaccess-berlin/berlin_declaration.pdf> Dichiarazione di Berlino. Accesso aperto alla letteratura scientifica, 2003, tr.it.Mornatti Susanna e Gargiulo Paola: <www.zim.mpg.de/openaccess-berlin/BerlinDeclaration_it.pdf>

Best Practices for Publishing Linked Data: <http://www.w3.org/2011/gld/wiki/Best_Practices_Discussion_Summary; <https://dvc.w3.org/hg/gld/raw-file/default/bp/index.html>>

Bethesda Statement on Open Access Publishing, 2003: <www.earlham.edu/~peters/fos/bethesda.htm>

Biblioteconomia per argomenti: <<http://www.frigimelica.it/vrd/argom.htm#dig>>

BIBO (Bibliographic Ontology): <<http://purl.org/ontology/bibo/>>

Blog “240inpratica”: <<http://240inpratica.net/accesso-aperto-e-statuti-universitari>>

Blog di R. Poynder: <<http://poynder.blogspot.ca/2013/06/open-access-springer-tightens-rules-on.html>> (Un tratto dal Blog: a gennaio 2013 Springer ha inasprito la sua OA self-archiving Policy imponendo un embargo di 12 mesi).

Blog *Scholarly Open Access: Critical Analysis of scholarly Open-Access publishing*: <<http://scholarlyoa.com/>>

Bollettino telematico di Filosofia Politica: <<http://bfp.sp.unipi.it/btftp/>>

Breve Introduzione all'Accesso Aperto, di Peter Suber. Online dal 29 dicembre 2004. Trad. it. di Susanna Mornati: <<http://www.aepic.it/docs/OA/brief-italian.htm>>

Budapest Open Access Initiative, 2002, (BOAI): <<http://www.soros.org/openaccess/index>>

Brisbane Declaration on Open Access and Research Conference, 2008, <<http://www.earlham.edu/~peters/fos/2008/10/brisbane-declaration-on-oa.html>>

Building Institutional Repositories for Global Research Commons. Sul Blog dell'AIMS (Agricultural Information Management Standards), 07/01/2013: <<http://aims.fao.org/community/open-access/blogs/building-institutional-repositories-global-research-commons>>

Cambridge University Press: <<http://www.cambridge.org/>>
CERIF (Common European Research Information Format): <<http://www.eurocris.org/Index.php?page=CERIF2008&t=1>>

CiTO, the Citation Typing Ontology, 2013: <<http://www.essepuntato.it/lode/http://purl.org/spar/cito>>

Classification and Visualization: Interfaces to Knowledge, International UDC Seminar, 24 - 25 October 2013, <<http://seminar.udcc.org/2013/programme.php>>

Compatibilità delle piattaforme DSpace, EPrints e quelle proprietarie con l'OpenAIRE: <<http://www.openaire.eu/it/component/content/article/9-news-events/340-how-to-make-your-repository-openaire-compliant-the-sessions-recordings-and-slides>>

Confederation of Open Access Repositories (CPAR). Task Force: <<http://www.coar-repositories.org/working-groups/repository-content/licenses-task-force/>>

Controlled Experiment for Reviewing Lifecycle Services. VOA3R Open Access Repository Project: <<http://voa3rexperiment12.wordpress.com/>>

Core Scientific Meta-Data (CSMD): <<http://code.google.com/p/icatproject/wiki/CSMD>>

Criteria for successful Repositories, in “A Scientist and the Web BLOG”, 2011:
<<http://blogs.ch.cam.ac.uk/pmr/2011/08/19/criteria-for-successful-repositories/>>

Cultural, Artistic and Scientific knowledge for Preservation, Access and Retrieval (CASPAR):
CDSware (CERN Document Server Software specializzato in fisica nucleare):
<<http://www.markosweb.com/www/cdsware.cern.ch/>>

CERIF-XML Data Exchange Format:
<http://www.eurocris.org/Uploads/Web%20pages/CERIF2008/Release_1.2/CERIF2008_1.2_XML.pdf>

CINECA - Consorzio Interuniversitario per il Calcolo Automatico dell'Italia Nord Orientale:
<<http://dirittoautore.cab.unipd.it/progetti/cineca-consorzio-interuniversitario-per-il-calcolo-automatico-dellitalia-nord-orientale>>

CKAN (Comprehensive Knowledge Archive Network) Registry: <<http://ckan.org/>>

Classificazione Decimale Dewey pubblicata come Linked Data (“dewey.info”):
<<http://dewey.info/>>

Classification & Visualization: Interfaces to Knowledge”, CFP: International UDC Seminar
2013, The Hague, 24-25 October, <http://universaldecimalclassification.blogspot.co.at/>

Clausola-modello per la Politica Open Access degli Atenei (Italia):
<http://240inpratica.net/accesso-aperto-e-statuti-universitari/>

COAR Repositories, news: <<http://www.coar-repositories.org/news-media/eurocris-and-coar-join-forces-building-up-a-mutual-partnership-2/>>

Comunicazione della Commissione al Parlamento Europeo, al Consiglio e al
Comitato Economico e Sociale Europeo sull'informazione scientifica nell'era digitale: accesso,
diffusione e conservazione, Bruxelles, 14.2.2007 COM(2007) 56 definitivo:
<<http://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2007:0056:FIN:IT:PDF>>

Confederazione di Open Access Repositories (COAR): <<http://www.coar-repositories.org/>>

Conferenza di Berlino: <<http://oa.mpg.de/openaccessberlin/signatories-extended.html>>

Creative Commons/CC: <http://it.wikipedia.org/wiki/Creative_Commons>

Current State of Open Access Repository Interoperability, COAR(Confederation of *Open Access*
Repositories), 2012: <<http://www.coar-repositories.org/files/COAR-Current-State-of-Open-Access-Repository-Interoperability-26-10-2012.pdf>>

DATABASE CITAZIONALI:

- Spires: <<http://www-library.desy.de/bluspires/hep/references.html>>;
- Citebase: <<http://users.ecs.soton.ac.uk/aw01r/citebase/evalForm1.htm>>, che analizza le citazioni degli e agli articoli della comunità dei fisici delle alte energie contenuti nell'archivio disciplinare in Arxiv <<http://arxiv.org/>>;

- CitEc: <<http://ideas.repec.org/top/top.series.recurse.html>> un proprio modello di misurazione delle citazioni che calcola l'impatto dei lavori di ricerca in economia contenuti nell'Archivio disciplinare RepEc

- Citeulike: <<http://www.citeulike.org>> favorisce la creazione di bibliografie condivise e interconnesse;

- Reference Linking: <<http://opcit.eprints.org/>> e

- Citation Analysis: <http://eprints.ecs.soton.ac.uk/10000/1/tim_oa.pdf>

che combinati assieme compongono un punto di partenza fondamentale entro il quadro *Open Archives Initiative* per i nuovi modelli di comunicazione scientifica in rete.

Databib: Cataloging the World's Research Data Repositories: an online tool for helping people identify and locate online REPOSITORIES of RESEARCH DATA:

<<http://databib.org/index.php>>

Data Exchange Agreement (DEA) per la trasmissione dei dati nella Biblioteca Digitale Europea: <<http://version1.europeana.eu/web/europeana/project/newagreement/>>

Data.bnf.fr: an overall presentation: *Linked Data* nella Biblioteca Nazionale di Francia : <<http://data.bnf.fr/docs/databnf-presentation-en.pdf>>

Data Manager Online, il Portale dell'ICT professionale:

<<http://www.datamanager.it/rivista/business-intelligence/la-data-governance-alla-base-degli-strumenti-di-bi>>

Data Management Plans. DCC: Digital Curation Center: <<http://www.dcc.ac.uk/resources/data-management-plans>>

DATI APERTI. Alcune prime indicazioni sull'assolvimento degli obblighi di trasmissione delle informazioni all'Autorità per la vigilanza sui contratti pubblici di lavori, servizi e forniture sulla base dell'art. 1, comma 32 della legge 190/2012:

- Deliberazione n. 26 del 22 maggio 2013 (pubblicata nella Gazzetta ufficiale n. 134 del 10/06/2013):

- <http://www.avcp.it/portal/public/classic/AttivitaAutorita/AttiDellAutorita/_Atto?ca=5396>

- Comunicato del Presidente del 22 maggio 2013:

- <http://www.avcp.it/portal/public/classic/AttivitaAutorita/AttiDellAutorita/_Atto?ca=5397>

- Le indicazioni operative e le specifiche tecniche per la pubblicazione dei dati:

- <<http://www.avcp.it/portal/rest/jcr/repository/collaboration/Digital%20Assets/pdf/AICom27.05.13SpecificeTecnichev1.0.pdf>>

Dati.camera.it – un portale italiano basato sui meccanismi LINKED DATA: <<http://dati.camera.it/it/linked-data/>>

Datos enlazados en la BNE, Biblioteca Nacional de España:

<<http://www.bne.es/es/Inicio/Perfiles/Bibliotecarios/DatosEnlazados/index.html>>

Datos.Bne.Es: <<http://datahub.io/dataset/datos-bne-es>>

Department for International Development “DFID Open and Enhanced Access Policy”,
<<https://www.gov.uk/government/publications/dfid-research-open-and-enhanced-access-policy>>

DepositoLegale.it: <<http://www.depositolegale.it/national-bibliography-number/>>

Dichiarazione di Principi internazionali di Catalogazione.IFLA , 2009:
<http://www.ifla.org/files/assets/cataloguing/icp/icp_2009-it.pdf>

“Diritto d’autore”, pubblicato sul PLEADI (Portale per la Letteratura scientifica Elettronica Italiana su Archivi aperti e Depositi Istituzionali): <<http://www.openarchives.it/pleiadi/open-access/diritto-dautore>>

"Diritto d'autore", pubblicato sul Sistema Bibliotecario dell'Università degli Studi di Padova:
<<http://www.cab.unipd.it/node/1104>>

DINI, la certificazione degli Archivi digitali: <<http://www.dini.de/english/dini-certificate/>;
<<http://edoc.hu-berlin.de/series/dini-schriften/2010-3-en/PDF/dini-zertifikat-2010-3-en.pdf>>

Directory of Open Access Repositories (DOAR): <http://www.opendoar.org/find.php>

Direttiva Europea 2013/37/UE sulle norme di utilizzo del patrimonio informativo del settore pubblico, intese a rafforzare l’accesso alle informazioni degli enti pubblici anche con l’utilizzo di nuove applicazioni: <<http://www.unipd.it/ilbo/content/un-formato-europeo-i-dati-di-biblioteche-musei-e-archivi>>

Dichiarazione di Berlino sull'accesso aperto alla conoscenza nelle discipline scientifiche e umanistiche, 2003: <<http://hdl.handle.net/10760/4585>>

Dichiarazione di Alhambra sull’ Accesso Aperto:
<http://oaseminar.fecyt.es/Resources/Documentos/ADeclaration/AD_OA_Italian.pdf>

Digital Libraries: Metadata Resources, IFLA: <<http://archive.ifla.org/II/metadata.htm>>

Diritto d’autore_Università di Padova: <<http://www.cab.unipd.it/servizi/diritto-dautore>>

DOAB (Directory of Open Access Books): <www.doabooks.org>

DOAJ (Directory of Open Access Journals): <<http://www.doaj.org/>>

DOAR (Directory of Open Access Repositories): <<http://www.opendoar.org/>>

DRIVER II: Digital Repositories Infrastructure Vision for European Research II:
<<http://www.madgik.di.uoa.gr/content/214>>

DRUPAL. Free and open-source Content Management System:
<<http://en.wikipedia.org/wiki/Drupal>>

Dublin Core® Metadata Initiative:

- DCMI Metadata Terms: <<http://dublincore.org/documents/dcmi-terms/>>;
- Dublin Core Abstract Model: <<http://dublincore.org/documents/abstract-model/>>

- Dublin Core Metadata Element Set (DCMES), Version 1.1, DCMI Recommendation, 2010-10-11, <<http://dublincore.org/documents/dces/>>; <<http://dublincore.org/documents/dces/#ISO15836>>
- Dublin Core Metadata Element Set - versione 1.1: traduzione italiana. ICCU: <<http://www.iccu.sbn.it/dublinco.html>>;
- Tools and Software (Conversion; Crosswalks; Metadata Creation, extraction, generation, harvesting, quality, translation, validation): <<http://dublincore.org/tools/>>;
- Dublin Core - Library Application Profile: <<http://dublincore.org/documents/library-application-profile/>>
- Dublin Core Metadata Template: <<http://www.ad.jyu.fi/users/a/ankarjal/ITKD50/Dublin%20Core%20Metadata%20Template.htm>>
- Dublin Core Qualifiers. DCMI Recommended Resource: <<http://dublincore.org/documents/usageguide/qualifiers.shtml>>
- Dublin Core User Guide/Publishing Metadata as Linked Data. Questa Guida offre una ricca spiegazione su come attribuire i valori “letterali” e “non-letterali” ai metadati DC: <http://wiki.dublincore.org/index.php/User_Guide/Publishing_Metadata>
- User Guide/Creating Metadata: <http://wiki.dublincore.org/index.php/User_Guide/Creating_Metadata#Guidelines_for_the_creation_of_medium_content>

Earth-prints: Archivio Disciplinare per le ricerche scientifiche in Geofisica, Vulcanologia, Scienze della Terra: <<http://www.earth-prints.org/>>

Editori OA per monografie: <http://oad.simmons.edu/oadwiki/Main_Page>

E-LIS. Archivio e-prints in Library and Information Science: <<http://eprints.rclis.org/>>

E-science: verso un network italiano per l'Open Access. Convegno, 25 ottobre 2013, CNR di Roma: <http://wiki.openarchives.it/index.php/Open_access_Week_-_2013>

Enabling Open Scholarship (EOS): <http://www.openscholarship.org/jcms/c_6162/en/repositories>

Enhanced publication: <<http://www.openaire.eu/en/component/content/article/76-highlights/344-a-short-introduction-to-enhanced-publications>>

euroCRIS and COAR join forces building up a mutual partnership: <<http://www.coar-repositories.org/news-media/eurocris-and-coar-join-forces-building-up-a-mutual-partnership-2/>>

EuroCRIS: <<http://www.eurocris.org/>>

EuroCRIS_Presentation: <http://www.vivoweb.org/files/presentations/12Fri/euroCRIS_LOD_and_%20VIVO.pdf>

European Commission background note on Open Access to publications and data in Horizon 2020, 2012: <http://ec.europa.eu/research/science-society/document_library/pdf_06/background-paper-open-access-october-2012_en.pdf>

European Research Council Open Access Policy: <<http://erc.europa.eu/about-erc/mission>>

European Interoperability Framework:<<http://www.foxesandfishes.it/wp/?p=657>>

EUROPEANA:< <http://www.europeana.eu/>>

EDM (Europeana Data Model): <<http://pro.europeana.eu/edm-documentation>>

FAST (Faceted Application of Subject Terminology): <<http://fast.oclc.org>>. Il servizio sperimentale di OCLC per i vocabolari controllati: <<http://tspilot.oclc.org/resources>>

Featured creators and stories: a facemeeting on open source for libraries. Conferenza per esaminare le nuove tendenze dei tools per la gestione delle biblioteche e la fruizione dei loro dati, la Pontificia Università della Santa Croce, l'American University of Rome, 1 ottobre 2013, <<http://www.pusc.it/bib/featured-creators-and-stories>>

FRAD - Functional Requirements for Authority:

<http://www.ifla.org/files/assets/cataloguing/frad/frad_2013.pdf>

(link alle traduzioni disponibili: <<http://www.ifla.org/publications/functional-requirements-for-authority-data>>)

Firenze University Press: <<http://www.fupress.com/index.asp>>

FOAF (Friend Of A Friend Ontology): <<http://xmlns.com/foaf/0.1/>>

Functional Requirements for Bibliographic Records (FRBR), vol. 19 of UBCIM, 1998 (Last update: 18 August 2011): <<http://www.ifla.org/publications/functional-requirements-for-bibliographic-records>>

Funzionalità offerte dall'Archivio Istituzionale dell'Università di Padova PHAIDRA:

<https://phaidra.cab.unipd.it/help_long#printed-guides>

G-8 International Conference on Open Data for Agriculture. Un'intervista con Johannes Keizer, in "Washington, D.C"., 29-30 April, 2013: <<http://aims.fao.org/community/agris-network/blogs/challenge-not-make-data-theoretically-accessible-also-practically-usab>>

German National Library goes LOD & publishes National Bibliography, Open bibliography and Open Bibliographic Data: <<http://openbiblio.net/2012/01/26/german-national-library-goes-lod-publishes-national-bibliography/>>

Guardian.co.uk: "Results of publicly funded research will be Open Access – science minister", Thursday 8 December 2011: <<http://www.guardian.co.uk/science/2011/dec/08/publicly-funded-research-open-access>>

GeoNames:<<http://www.geonames.org/>>

Ghent Declaration:

<http://www.openaire.eu/index.php?option=com_content&view=article&id=223%3Aseizing-the-opportunity-for-open-access-to-european-research-ghent-declaration-published&catid=76%3Ahighlights&lang=it>

Global Interoperability and Linked Data in Libraries. Presentazioni del seminario internazionale, Firenze, June 2012, in “Jlis.it”, 3 (2012):

<<http://leo.cilea.it/index.php/jlis/issue/view/368/showToc>>

Glossario delle organizzazioni internazionali e dei termini più comunemente utilizzati negli Open Archives: <<http://www.biblioteche.unitus.it/download/openaccess/risorse/glossario.pdf>>

Glossary of Metadata Standards, Indiana University Libraries White Professional Development Award, 2010:

<http://www.dlib.indiana.edu/~jenrile/metadatamap/seeingstandards_glossary_pamphlet.pdf>

Google Book: digitalizzazione di massa (una class action contro Google):

<<http://www.unipd.it/ilbo/content/nome-del-copyright-una-class-action-contro-google-no-grazie>>

Government news G8: <<https://www.gov.uk/government/news/g8-science-ministers-statement>>

Gratis and Libre Open Access”, Peter Suber, 2008:

<<http://legacy.earlham.edu/~peters/fof/newsletter/08-02-08.htm>>

Guida pratica all’auto-archiviazione “Come posso auto-archiviare e pubblicare il mio lavoro?”, trad. it. di How can I self-archive AND get my paper published?:

<<http://paduaresearch.cab.unipd.it/docs/romeo-it-posso-archiviare.pdf>>

Guidelines for Dublin Core Application Profile, 2009: <<http://dublincore.org/documents/profile-guidelines/>>

“Guida “Good practices for university Open-Access Policies” fornita dal progetto Harvard Open Access Project (HOAP):

<http://cyber.law.harvard.edu/hoap/Good_practices_for_university_open-access_policies>

Guide to Institutional Repository Software. A cura dell’Open Society Institute, contiene un’analisi dei principali software per creare un Digital Repository:

<http://www.cab.unipd.it/system/files/OSI_Guide_to_IR_Software_v3.pdf>

Horizon 2020 Italia (HIT 2020) - piano programmatico per la ricerca in Italia:

<<http://hubmiur.pubblica.istruzione.it/web/ministero/focus190313>>;

<<https://www.researchitaly.it/conoscere/stampa-e-media/news/horizon-2020-italia/>>

How to Cite Datasets and Link to Publications, Digital Curation Centre:

<<http://www.dcc.ac.uk/resources/how-guides/cite-datasets>>

How to facilitate the linking between resources? AIMS (Agricultural Information Management Standards), 2012: <<http://aims.fao.org/linked-data/getting-started>>

Hungarian National Library OPAC and Digital Library Published as Linked Data, 2010,

ISKO UK: <<http://iskouk.blogspot.com/2010/05/hungarian-national-library-opac-and.html>>

Hypothèses.org. Portale per il blogging in ambito accademico (Scholarly blogging in the Social Sciences and Humanities): <<http://blogs.eui.eu/library/hypotheses-org-scholarly-blogging-in-the-social-sciences-and-humanities.html>>

I Profili di Applicazione: <<http://www.iccu.sbn.it/upload/documenti/Peruginelli.pdf>>

IA Bookreader: il lettore di libri Open Source: <<http://openlibrary.org/dev/docs/bookreader>>

IFLA, Semantic Web Special Interest Group: <<http://www.ifla.org/swsig>>

Il lato oscuro dell'*Open Access*, in “La Repubblica”, 04 ottobre 2013:
<http://www.lescienze.it/news/2013/10/04/news/riviste_open_access_mancato_peer_review-1835164/>

Incentives, Integration, and Mediation: Sustainable Practices for Populating Repositories. Pubblicato da COAR (Confederation of Open Access Repositories), giugno, 2013:
<http://www.coar-repositories.org/files/Sustainable-best-practices_final.pdf>

In defence of *Open Access* systems, su Blog “The Hindu”, December 31, 2012:
<<http://www.thehindu.com/features/education/research/in-defence-of-open-access-systems/article4258948.ece>>

Insight Discoverer™ Clusterer:
<http://registry.dfki.de/sections.php3?f_mainsection=2&f_section=50&f_system=414>

Institutional Repositories for Research Management and Assessment, on the “Open Access scholarly Information Sourcebook” Portal: <http://openoasis.org/index.php?option=com_content&view=article&id=165&Itemid=335>

ISO ISO:14721 OAIS (*Open Archival Information System*);

ISO 16363:2012 --Space data and information transfer systems -- Audit and CERTIFICATION of TRUSTWORTHY DIGITAL REPOSITORIES (TDR):
<http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=56510>

ISO 26324:2012 -- Information and documentation -- DIGITAL OBJECT IDENTIFIER (DOI) system:
<http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=43506>

JISC (Joint Information Systems Committee): “Digital Preservation briefing paper”, 2010:
<http://www.jisc.ac.uk/publications/briefingpapers/2006/pub_digipreservationbp.aspx>

JSTOR: <<http://www.jstor.org/>>

L'equivalenza nella terminologia multilingue, Breve introduzione alla terminologia:
<http://www.farum.it/intro_terminologia/ezine_articles.php?id=25>

Labordoc. Institutional Repository (con le risorse digitalizzati a partire dal 1919): <http://labordoc.ilo.org/>

L'AperLA, progetto OCR per il recupero dei testi in cattivo stato di conservazione dell'Istituto di Linguistica Computazionale: <http://www.ilc.cnr.it/viewpage.php/sez=ricerca/id=75/vers=ita>

Learning Linked Data project: <http://lld.ischool.uw.edu/wp/About/>

LEDIZIONI è la prima casa editrice commerciale italiana a rendere disponibile le sue monografie Open Access su DOAB (Directory of Open Access Books): <http://www.doabooks.org/doab?func=publisher&pId=1201&uiLanguage=>

LEGGE 7 ottobre 2013, n. 112. Conversione in legge, con modificazioni, del decreto-legge 8 agosto 2013, n. 91, recante Disposizioni Urgenti per la Tutela, la Valorizzazione e il Rilancio dei beni e delle attività culturali e del turismo (13G00158) (GU n.236 del 8-10-2013): <http://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:legge:2013;112>

LIBER (Ligue des Bibliothèques Européennes de Recherche – Association of European Research Libraries): <http://www.libereurope.eu/>

Libraries and the Semantic Web. Session 149 — Information Technology, Cataloguing, Classification and Indexing with Knowledge Management, 15 August 2010: <http://www.ifla.org/node/7177>

LIBRARY LINKED DATA *Incubator Group*: <http://www.w3.org/2005/Incubator/lld/>
<http://thedatahub.org/group/lld>

LIBRARY LINKED DATA *Incubator Group* Final Report:
<http://www.w3.org/2005/Incubator/lld/XGR-lld-20111025/>

LIBRARY LINKED DATA *Incubator Group*: Datasets, Value Vocabularies:
http://www.w3.org/2005/Incubator/lld/wiki/Vocabulary_and_Dataset

LIBRARY OF CONGRESS (LC) Authorities: <http://authorities.loc.gov/>; LC Linked Data Service Authorities and Vocabularies: <http://id.loc.gov/>

LIBRARY SCIENCE TALK. Presentazione dei lavori di LIBRARY LINKED DATA Incubator Group, 12-13 marzo 2012:
http://www.nb.admin.ch/aktuelles/ausstellungen_und_veranstaltungen/00726/01611/03953/03958/index.html?lang=en

LIBRE' OPEN ACCESS: <http://bit.ly/libre-oa>

LICENZE:

- con accordi aggiuntivi Addendum (traduzione in italiano): <http://paduaresearch.cab.unipd.it/docs/SPARC%20AUTHOR%20ADDENDUM%20traduzione.pdf>
- formulate da SPARC: <http://www.sparc.arl.org/>
- [Science Commons](http://creativecommons.org/science) <http://creativecommons.org/science>
- di JISC <http://www.jisc.ac.uk/> e SURF <http://www.surf.nl/en/Pages/default.aspx>

costituiscono soluzioni valide e già applicate in molti paesi.

LINEE GUIDA:

- MedOANet *for implementing OPEN ACCESS POLICIES*, 2013:
<<http://www.medoanet.eu/news/open-access-week-medoanet-project-releases-guidelines-implementing-open-access-policies-research>>
- per l'Interoperabilità Semantica attraverso i LINKED OPEN DATA, Sistema Pubblico di connettività e cooperazione, 2012:
<http://www.digitpa.gov.it/sites/default/files/CdC-SPC-GdL6-InteroperabilitaSemOpenData_0.pdf>
- DRIVER (*Digital Repository Infrastructure Vision for European Research*):
<<http://www.driver-repository.eu/>>
- OpenAIRE (*Open Access Infrastructure Research for Europe*):
<<http://www.openaire.eu/component/content/article/207>>
- CRUI per gli Archivi Istituzionali, 2009:
<<http://www.cruis.it/HomePage.aspx?ref=1781>>
- CRUI per le Riviste elettroniche ad Accesso Aperto, 2009:
<<http://www.cruis.it/HomePage.aspx?ref=1789>>
- CRUI per il deposito delle Tesi di Dottorato negli Archivi Aperti, 2007:
<<http://www.cruis.it/HomePage.aspx?ref=1149>>
- CRUI per la creazione e la gestione di metadati nei Repository Istituzionali, 2012:
<<http://www.cruis.it/HomePage.aspx?ref=2066>>
- per "I Diritti dell'Autore" redatte da Scholarly Publishing and Academic Resources Coalition: <http://paduaresearch.cab.unipd.it/docs/SPARC_AUTHORITYRIGHTS2007-it.pdf> <<http://www.unimi.it/ricerca/air/20413.htm>>

LINKED DATA API and Format - to simplify use of LINKED DATA by web – developers:
<https://code.google.com/p/linked-data-api/wiki/API_Vocabulary>

LINKED DATA *Glossary* (Author: Bernadette Hyland), work in progress, 2013,
<<http://vocabulary.semantic-web.at/w3c-linked-data-glossary>> (È un Glossario di termini definiti e usati per descrivere Linked Data, vocabolari associati e Best Practices. È pubblicato tramite i formalismi SKOS (Simple Knowledge Organization Systems) come risorsa LOD (Linked Open Data) + SPARQL endpoint da W3C)

LINKED DATA *for Libraries: short introduction to the concepts and technology behind Linked Data, how it works, and some benefits it brings to libraries*, OCLC video, 2013:
<<http://www.youtube.com/watch?v=fWfEYcnk8Z8>>
<<http://www.youtube.com/watch?v=fWfEYcnk8Z8x%20x%20>>

LINKED OPEN DATA. *What is it?* (from Europeana), video, 2013:
<<http://vimeo.com/36752317>>

LINKED DATA for Holdings and Cataloguing: The First Step Is Always the Hardest! , in "An introductory overview of Linked Open Data in the context of cultural institutions", AUDIO & SLIDES: The sessions Linked Data for Holdings and Cataloguing by Eric Miller (Zepheira) and Richard Wallis (OCLC), 2013:
<<http://eventscribe.com/2013/ALA-Midwinter/aaSearchByDay.asp>>

LINKED OPEN DATA-enabled *Bibliographical Data Recommendations*, LOD-BD, AIMS, 2012: <<http://www.digitalmeetsculture.net/wp-content/uploads/2012/10/LODE-BD-2.pdf>>

LODE-BD: 3.7. Rights. Situations and best practices for encoding the data related to "Rights", "Terms of use", and "Access condition": <<http://aims.fao.org/lode/bd-2/rights>>
LODE-BD. Appendix 4. Crosswalk of Metadata Terms used in LOD-BD and Schema.org Terms: <<http://aims.fao.org/lode/bd-2/schema-org-crosswalk>>

LINKED DATA *Prototyping* at the British Library:
<<http://talis-linkeddata-libraries.s3.amazonaws.com/Linked%20Data%20Prototyping.pdf>>

LINKED DATA, *Connect Distributed data across the web* (the site of the Linked Data community): <<http://linkeddata.org/>>

LINKED DATA on the Web (LDOW2014), 7th Workshop, 23rd International World Wide Web Conference, 7-11 April 2014, Seoul, Korea: <<http://events.linkeddata.org/ldow2014/>>

LINKED OPEN DATA Italia. Dati aperti collegati e usabili: <<http://www.linkedopendata.it/>>

LINKED OPEN DATA Pilot description (Europeana): <<http://pro.europeana.eu/linked-open-data>>

LINKED OPEN DATA Conference (2nd)from the Cataloguing and Indexing Group in Scotland (CIGS). *Opening Library Linked Data to National Heritage: Perspectives on International Practice*, September 2012: <<http://aims.fao.org/events/2nd-linked-open-data-conference-cataloguing-and-indexing-group-scotland-cigs-opening-library->>

LINKING to the Future, DC-2013 conference: <<http://dublincore.org>>

LINKED, OPEN, SEMANTIC? Paul Walk's weblog, 2009:
<<http://blog.paulwalk.net/2009/11/11/linked-open-semantic/>>

LINKED OPEN VOCABULARIES (LOV) Service: un dataset di descrizioni dei vocabolari o ontologie del Web Semantico nei linguaggi RDFS o OWL definiti per o usati da dataset Linked Data:
<<http://labs.mondeca.com/dataset/lov/index.html>>; <<http://lov.okfn.org/dataset/lov/index.html>>

LINKING OPEN DATA *Cloud Diagram*: <<http://richard.cyganiak.de/2007/10/lod>>; CKAN: Linking Open Data Cloud: <<http://datahub.io/group/lodcloud>>

LOD-Around-The-Clock (LATC), EC FP7 Support Action: <<http://5stardata.info>>

LOD2Project: Creating qualitative Knowledge out of InterLINKED Data:
<<http://lod2.eu/WikiArticle/Project.htm>>; <<http://lod2.eu/Welcome.html>>

LOD - LOBID. ORG. Il servizio *Linked Open Data* della Biblioteca di North Rhine-Westphalia:
<<http://lobid.org/about>>

LODLAM, *Linked Open Data in Libraries, Archives, and Museums*, <<http://lodlam.net/>>

MACS (Multilingual Access to Subject), un Authority File virtuale multilingue:

<http://www.sba.unifi.it/ac/relazioni/clavel-merrin_ita.pdf>

MAG: Metadati Amministrativi e Gestionali. Manuale Utente, versione: 2.0.1, ICCU, 2006,

<<http://www.iccu.sbn.it/opencms/opencms/documenti/manuale.html>>;

<<http://www.iccu.sbn.it/upload/documenti/Manuale.pdf>>

Manuale per l'interazione con gli utenti del Web culturale:

<http://www.minervaeurope.org/publications/handbookwebusers_it/chapter4_5.html>

MARiMbA: <<http://mayor2.dia.fi.upm.es/oeg-upm/index.php/en/downloads/228-marimba>>

MELIBEA: <<http://www.accesoabierto.net/politicas/default.php>>

Messina Declaration, 2004: <<http://eprints-phd.biblio.unitn.it/help/MessinaEN.pdf>>;

<<http://www.aepic.it/conf/Messina041/index981f.html>>

Metadata & Semantics for Open Access Repositories, Research Information Systems and Infrastructures, 6th Metadata and Semantics Research Conference, 28-30 November 2012 -

Cádiz (Spain): <http://mtsr2012.uca.es/track_open_access.php>; The 5th International Conference on Metadata and Semantics Research (MTSR'11) :

<<http://ieru.org/org/mtsr2011/OAR/>>

Metadata Management in OCLC (Online Computer Library Center):

<<http://www.oclc.org/research/activities/msm/about.html>>

Metadata Mapping between metadata formats, UKOLN: The UK Office for Library and Information Networking, University of Bath:

<<http://www.ukoln.ac.uk/metadata/interoperability/>>

MedOANet (Mediterranean Open Access Network), un progetto europeo volto al miglioramento di strategie, politiche e strutture per l'accesso aperto in sei Paesi del Mediterraneo: Italia, Grecia, Francia, Spagna, Portogallo e Turchia: <<http://www.medoanet.eu/project-it>>

MetaMap: <<http://www.mapageweb.umontreal.ca/turner/meta/english/index.html>>

Metadata for Public Sector Administration, Asset Description Metadata Schema, 2013:

<<http://www.niso.org/news/events/2013/dcmi/publicsector/>>;

<<https://joinup.ec.europa.eu/asset/adms/release/100>>

Metadata and Semantics for Agriculture, Food and Environment. Workshop Internazionale MTSR 2012:<<http://eco.logismi.co/events/mtsr2012/>>

METS (Metadata Encoding and Transmission Standard): <http://www.loc.gov/standards/mets/>

Miriam-Webster online: <<http://www.merriam-webster.com/dictionary/schema>>

MIT OpenCourseWare: <http://ocw.mit.edu/>

MODELLI ECONOMICI di finanziamento del materiale OPEN ACCESS:

- “author pays”: <<http://www.nature.com/nature/focus/accessdebate/26.html>>
- finanziamenti delle istituzioni, modello BioMed Central: <<http://www.biomedcentral.com/info/about/instmembership>>;
- oppure iniziative editoriali autofinanziate: Società Scientifiche, University Press: <<http://www.arl.org/sparc/GI/>>;
- Economic Implications of Alternative Scholarly Publishing Models: Exploring the costs and benefits, JISC 2009: <<http://www.jisc.ac.uk/publications/reports/2009/economicpublishingmodelsfinalreport.aspx>>
- "Pubblicazioni ad Accesso Aperto: Modelli Economici alternativi, sostenibilità e vantaggi per i fruitori". Relazioni dal seminario "Il Sistema Bibliotecario dell'Università del Piemonte Orientale", 28.10.2011: <<http://www.unipmn.it/Sistema%20Bibliotecario%20di%20Ateneo/Open%20Access/Open%20access/default.aspx>>
- OA Book business models lanciato da una comunità Wiki “The Open Access Directory (OAD)” : <http://oad.simmons.edu/oadwiki/OA_book_business_models>

MOTORI di RICERCA del Materiale Scientifico:

- BASE (Bielefeld Academic Search Engine): <<http://www.base-search.net/>>
- OAIster: <<http://www.oclc.org/oaister.en.html>>
- Open Access Thesis and Dissertations: <<http://www.oatd.org/>>
- PubPsych, un aggregatore dei metadati nell'ambito della psicologia con particolare attenzione alla produzione scientifica europea. PubPsych aggrega i metadati provenienti dai principali repertori online, repository e portali di ricerca in ambito psicologico i cui metadati sono liberamente accessibili: <<http://pubpsych.zpid.de/pubpsych/>>

Munin Conference on Scientific Publishing:

<<http://webtv.uit.no/Mediasite/Catalog/Full/a6f8a77feefb4eaa802ef0e701338ece21>>

NACO: Library of Congress Authorities: <<http://authorities.loc.gov/>>

NANOPUB.ORG: <http://nanopub.org/wordpress/?page_id=65>

National Information Standards Organization (NISO): <<http://www.niso.org/home/>>

National Library of Australia: <<http://www.nla.gov.au/>>

National Open Access Desk: <<http://www.openaire.eu/it/open-access/country-information>>

National Repository of Open Education Resources. Le risorse sono disponibili sotto la licenza CC-BY-SA che permette la condivisione e l'uso commerciale delle risorse:

<<http://nroer.in/browserError/>>

National Science Digital Library (NSDL). Documentazione da Wiki per “Contributors and Developers”: <<http://nsdl.org/contribute>>

NDLTD (Networked Digital Library of Theses and Dissertations): <<http://www.ndltd.org/>>

Network nazionale di Archivi Digitali Istituzionali in Argentina: <<http://www.coar-repositories.org/news/argentina-develops-national-system-of-digital-repositories/>>

NISO Framework Working Group. A Framework of Guidance for Building Good Digital Collections. 3rd ed. Baltimore: NISO, 2007:
<<http://www.niso.org/publications/rp/framework3.pdf>>.

OAIS (Open Archival Information System), ISO (ISO 14721:2012):

<<http://parer.ibr.regione.emilia-romagna.it/notizie/on-line-la-nuova-versione-di-oais>>

OCLC provides downloadable Linked Data file - for the 1 million most widely held works in WorldCat:< <http://www.oclc.org/us/en/news/releases/2012/201252.html>>. (I contenuti in un unico file LD sono scaricabili da:

<<http://purl.oclc.org/dataset/WorldCat/datadumps/WorldCatMostHighlyHeld-2012-05-15.nt.gz>>)

ODIN Project. ORCID and DataCite Interoperability Network: Towards increasing the interoperability of PERSISTENT author and object IDENTIFIERS: <<http://odin-project.eu/mission/>>

OECD Principles and Guidelines for Access to Research Data from Public Funding: <<http://www.oecd.org/science/sci-tech/oecdprinciplesandguidelinesforaccesstoresearchdatafrompublicfunditions:ng.htm>>

OAI-ORE (Open Archives Initiative Object Reuse and Exchange): <<http://www.openarchives.org/ore/>>

OAIster: <<http://www.oclc.org/oaister/>>

Online Computer Library Center (OCLC): <<http://www.oclc.org/global/default.htm>>

Open Access (una bibliografia ben annotata):

<<http://scholarship20.blogspot.com/2012/10/resources-on-open-access-u-of-toronto.html>>

Open Access Books per la letteratura dell'area umanistica:

<<http://www.earlham.edu/~peters/fos/newsletter/08-02-06.htm#bookprojects>>

OPEN ACCESS BUTTON – a browser-based tool to map the epidemic of denied access to academic research articles, and help users find the research they need – with alternative access to the article, Berlin OA 11, 2013:

<<http://oabutton.wordpress.com/2013/11/18/students-launch-button-to-put-denied-access-to-research-on-the-map/>> www.OpenAccessButton.org.

OPEN ACCESS CLAUSES *in Publishers' Licenses: Current State and Lessons Learned*, by the Open Access Agreements and Licenses Task Force, COAR, October 24, 2013:

<<http://www.coar-repositories.org/files/OA-Clauses-in-Publishers-Licenses.pdf>>

Open Access reaches a turning point: Science-Metrix presents a study on Open Access to the European Commission's Directorate-General for Research and Innovation, Science Metric Inc., 2013: <http://www.science-metrix.com/eng/news_13_08.htm>

Open Access dall'IFLA (International Federation of Library Associations and Institutions):

<<http://www.ifla.org/strategic-plan/key-initiatives/digital-content/oa>>

Open Access e comunicazione scientifica, Università degli studi della Tuscia. Sistema Bibliotecario di Ateneo: <www.unitus.it/biblioteche/webif/06-docs/06-risorse/oa/oa01.html>

Open Access in Italia 2012:

<[http://wiki.openarchives.it/index.php/Dati e cifre sull%27Open Access in Italia - 2012](http://wiki.openarchives.it/index.php/Dati_e_cifre_sull%27Open_Access_in_Italia_-_2012)>

Open Access in Italia 2013:

<[http://wiki.openarchives.it/index.php/Open access Week - 2013](http://wiki.openarchives.it/index.php/Open_access_Week_-_2013)>

OA 2013 un elenco (semi) esaustivo degli eventi in programma in tutta Italia:

<[http://wiki.openarchives.it/index.php/Open access Week - 2013](http://wiki.openarchives.it/index.php/Open_access_Week_-_2013)>

Open Access: il filo arancione della ricerca scientifica, la Notte dei ricercatori, 27 settembre 2013:

<<http://www.unipmn.it/Sistema%20Bibliotecario%20di%20Ateneo/Open%20Access/Notte%20dei%20Ricercator/default.aspx>>

Open Access Map: <<http://www.openaccessmap.org/>>

Open Access Overview (definition, introduction)", Peter Suber: updated 2012:

<<http://legacy.earlham.edu/~peters/fos/overview.htm>>

Open Access Week: <<http://www.openaccessweek.org/>>

Open Access e condivisione delle conoscenze: l'impatto del modello nelle comunità di apprendimento. Programma di Ricerca 2005-2006. Principale ricercatore: Anna Maria Tammaro: <http://dspace-unipr.cilea.it/bitstream/1889/452/1/progetto_ricerca.pdf>

Open Access in the EU and in the European Research Area (Open access to scientific information), February 2013:

<http://www.consilium.europa.eu/ueDocs/cms_Data/docs/pressData/en/intm/135548.pdf>

Open Access Information and Resources, concept map created by Laura Briggs, 2010, <<http://robynhall.ca/openaccess/>>

Open Access CORE (COnnecting REpositories): <<http://core.kmi.open.ac.uk/search>>

OpenAIRE (Open Access Infrastructure for Research in Europe): <<https://www.openaire.eu/>>

OpenAIRE's 'Enhanced Publications' Pilots, 12 October 2012: <<http://www.openaire.eu/en/component/content/article/9-news-events/424-subject-specific-pilots-for-enhanced-publications>>

Open Archives Initiative (OAI): <<http://www.openarchives.org/data/registerasprovider.html>>

Open Archival Information System (OAIS) - Reference model. ISO 14721:2012: <<http://webstore.ansi.org/RecordDetail.aspx?sku=ISO+14721%3a2012#.UU7HyhxWySo>>

OpenAGRIS (web application to aggregate information from different Web sources with more than 60 million triples): <<http://aims.fao.org/openagris>>; <<http://agris.fao.org/agris-search/search/display.do?f=2010%2FJP%2FJP1014.xml%3BJP2010001379>>

OpenDOAR - Directory of Open Access Repositories: <<http://www.opendoar.org/>>

OpenDOAR_Help: <<http://opendoar.org/help.html#policies>>

Open Knowledge Foundation Italia - Sapere libero per tutti: <<http://it.okfn.org/>>

Open Science Federation: <<http://opensciencefederation.com/>>

OAI-PMH (Protocol for Metadata Harvesting):

<<http://www.openarchives.org/OAI/openarchivesprotocol.html>>

OpenAIRE fornisce Linee Guida, Buone Prassi, modelli e tecnologie per depositare, accedere a e manipolare i contenuti di ricerca di varie tipologie in combinazione con le pubblicazioni di riferimento: <<http://www.openaire.eu/it/open-access/mandates-a-policies>>

Open Book Publisher: <<http://www.openbookpublishers.com/>>; <<http://legacy.earlham.edu/~peters/fos/newsletter/08-02-08.htm>>

OpenDOAR (Directory of Open Access Repositories): <<http://www.opendoar.org/find.php>>

Open Annotation Community Group (OACG):

<<http://www.w3.org/community/openannotation/>>; (Annotation Ontology:

<<http://code.google.com/p/annotation-ontology/>>; Open Annotation Model:

<<http://www.openannotation.org/spec/beta/>>)

OAPEN-UK (2010-2015) per la pubblicazione delle monografie Open Access: <<http://oapen-uk.jiscebooks.org/>>

Object Constraint Language. Grammar: <<http://www.csci.csusb.edu/dick/samples/ocl.html>>

OpenAIRE: <<http://www.openaire.eu/it>>

OpenAIRE, “It's all in the Context...”, 14 February 2013: <<http://www.openaire.eu/en/component/content/article/76-highlights/443-its-all-in-the-context>>

OpenAIREplus. Aprendo la strada ad uno spazio aperto per l'informazione scientifica, 15 Dicembre 2011: <<http://www.openaire.eu/it/component/content/article/326-openaireplus-press-release>>

OpenAIRE Orphan Repository: <<http://openaire.cern.ch/>>

Open Archive Initiative: <<http://www.openarchives.org/>>

Open Archive Initiative – Protocol for Metadata Harvesting (OAI-PMH): <<http://www.openarchives.org/OAI/openarchivesprotocol.html>>

Open Data Charter and Technical Annex del G8 - Policy paper (Il principio 1 è: Open Data by default), 2013:

<<https://www.gov.uk/government/publications/open-data-charter/g8-open-data-charter-and-technical-annex>>

Open Data Handbook. Glossario, Open Knowledge Foundation:

<<http://opendatahandbook.org/it/glossary.html#term-application-programming-interface>>

Open definition. Open Knowledge Foundation Italia: <<http://it.okfn.org/>> (V. Licenza CC0 1.0 Universal CC0 1.0: <<http://creativecommons.org/publicdomain/zero/1.0/>>

Open Journal Systems: <<http://pkp.sfu.ca/?q=ojs>>

Open Library: <<http://openlibrary.org/>>

Open Repositories 2013. Repository Island, Charlottetown, PEI, Canada, <<http://or2013.net/program/session-schedule>>

Open-Researcher-Contributor ORCID: <<http://www.educause.edu/ero/article/open-researcher-contributor-id-orcid-solving-name-ambiguity-problem>>

OWL Web Ontology Language. Overview: <<http://www.w3.org/TR/owl-features/>>

“QScience.com” is one of the best collaborative, peer-reviewed online publishing platforms: <<http://www.qscience.com/page/about>>

Quinta edizione della Settimana internazionale dell'Open Access (dal 21 al 27 ottobre 2013) in Italia. Iniziative: <http://wiki.openarchives.it/index.php/Open_access_Week_-_2013>

- l'Università di Firenze:
“Il quadro legislativo italiano e la Politica dell'Open Access”:
<<http://www.sba.unifi.it/Article458.html>>;
FLORE – Florence Research Repository Istituzionale, pp. 6, 15:
<http://www.unifi.it/upload/sub/newsletter_unifi/newsletter_149.pdf>
- l'Università di Palermo: La presentazione di due pubblicazioni promozionali: “L'Open Access al servizio della ricerca”; “L'Open Access per una ricerca e un'innovazione mirate” <<http://portale.unipa.it/amministrazione/area1/ssp04/set11/RisorseOA/>>;
- l'Università del Piemonte Orientale. Ridefinire i diritti: autore e lettore nella prospettiva dell'Accesso Aperto alla letteratura scientifica. Le relazioni e il video:
<<http://www.unipmn.it/Sistema%20Bibliotecario%20di%20Ateneo/Open%20Access/Ridefinire%20i%20diritti/default.aspx>>;
- la Notte dei Ricercatori, il 27 settembre ad Alessandria. Il video delle scenette sull'*Open Access*:
<<http://www.unipmn.it/Sistema%20Bibliotecario%20di%20Ateneo/Open%20Access/Notte%20dei%20Ricercatori/default.aspx>>

PHAIDRA (Permanent Hosting, Archiving and Indexing of Digital Resources and Assets) – Archivio digitale - personalizzazione del Software Open Source Fedora: <<https://wiki.duraspace.org/display/FCCommReg/Phaidra>>

PHAIDRA – l'Università di Vienna: <<http://bibliothek.univie.ac.at/english/phaidra.html>>

PHAIDRA – l'Università di Padova: <<https://phaidra.cab.unipd.it/static/locandinaphaidra.pdf>>

PHAIDRA International: <<http://www.phaidra.org/>>

Pinakes Text (Strumento di filologia computazionale):

<<http://pinakes.imss.fi.it/index.php/Pinakestext>>

PLEADI": Archivi Istituzionali: <<http://www.openarchives.it/pleiadi/progetto-pleiadi/risorse-indicizzate>>

Policy Perspectives, To publish and perish, 1998:

<<http://www.arl.org/scomm/pew/pewrept.html>>

Policy sull'Open Access del Research Councils UK che armonizza le Policy già esistenti dal 2005 negli Enti di ricerca afferenti: <<http://www.rcuk.ac.uk/research/Pages/outputs.aspx>>

Politiche di Copyright: http://paduaresearch.cab.unipd.it/politiche_copyright.html

Politiche di gestione della conservazione delle collezioni digitali: Strategie a breve termine per contrastare problemi a lungo termine: <<http://www.dpworkshop.org/dpm-ita/terminology/repository.html>>

PORTALI a supporto e per la disseminazione dei contenuti di ricerca *Open Access*:

- COAR (Confederation of Open Access Repositories): <<http://www.coar-repositories.org/news-media/spanish-translation-of-the-open-access-policy-kit/>>
- DART-Europe Portal: <<http://www.dart-europe.eu/basic-search.php>>
- DRIVER (Digital Repository Infrastructure Vision for European Research) search portal: <<http://www.driver-repository.eu/>>
- EIFL (Knowledge without boundaries): <<http://www.eifl.net/who-we-are>>
- EnablingOpenScholarship (EOS): <http://www.openscholarship.org/jcms/j_6/accueil>
- NARCIS (National Academic Research and Collaborations Information System): <<http://onlinelibrary.maastrichtuniversity.nl/database/narcis/>>
- OpenAIRE (Open Access infrastructure for Research in Europe): <<http://www.openaire.eu/>>
- PLEADI (la piattaforma nazionale per l'accesso centralizzato alla letteratura scientifica depositata negli Archivi Aperti Italiani, istituiti e gestiti da Università ed Enti di ricerca): <http://wiki.openarchives.it/index.php/Pagina_principale>
- Public Library of Science (PLoS): <<http://www.plos.org/>>
- SPARC (Scholarly Publishing and Academic Resources Coalition): <<http://sparceurope.org/open-access/>>
- ResearchItaly (Portale della ricerca italiana): <<https://www.researchitaly.it/fare/ricerca-e-societa/comunicare/>>
- SPARC Europe: <<http://sparceurope.org/repositories/>>
- UNESCO's Global Open Access Portal (GOAP): <<http://www.unesco.org/new/en/communication-and-information/portals-and-platforms/goap/>>

Portale del Diritto: <[HTTP://IT.WIKIPEDIA.ORG/WIKI/PORTALE:DIRITTO](http://it.wikipedia.org/wiki/Portale:Diritto)>

Poseidon tool for UML: <<http://freecode.com/projects/poseidon>>

Position statement in support of Open and unrestricted Access to published research: <<http://www.wellcome.ac.uk/About-us/Policy/Policy-and-position-statements/WTD002766.htm>>

PRELIDA project. Preserving Linked Data: building bridges across the Digital Preservation and Linked Data communities, 2013 – 2015: <<http://www.cnr.it/cnr/news/CnrNews?IDn=2650>>

Preservation and Long-term Access through NETworked Services (PLANETS) Project: <<http://www.planets-project.eu/>>

Proposta di azione normativa sull'Accesso Aperto in Italia. Progetto europeo *MedOAnet* (*Mediaterranean Open Access Network*), 2012: <http://wiki.openarchives.it/index.php/Proposta_di_azione_normativa_sull%27accesso_aperto_in_Italia>

PubPsych: <<http://pubpsych.zpid.de/pubpsych/>>

(è un nuovo strumento appena rilasciato per accedere a risorse informative (articoli, libri, capitoli di libri, tesi di dottorato, test, dati della ricerca e trattamenti) in ambito psicologico con un focus sulla produzione scientifica europea liberamente accessibile).

Public access to publicly-funded research. Il discorso alla Publishers Association annual general meeting (London), Willetts, Hon David, 2 May 2012: <<http://www.bis.gov.uk/news/speeches/david-willetts-public-access-to-research>>

Public Library of Veroia in Web 3.0: <<http://gr.okfn.org/2012/10/libver/?lang=en>>

“Publication Open Repository Torino” (PORTO): <<http://porto.polito.it/>>

Pubblicazioni dei progetti finanziati dall'ERC:

<<http://www.openaire.eu/component/attachments/download/3en/component/attachments/download/3>>

RACCOMANDAZIONI:

- Budapest Open Access Initiative (BOAI-2012). Programma BOAI per i prossimi dieci anni: <<https://www.eff.org/deeplinks/2012/09/BOAI-10th-anniversary-new-recommendations>> <<http://www.soros.org/openaccess/boai-10-recommendations>>
- Raccomandazioni sull'Accesso all'informazione scientifica e sulla sua conservazione, 2012: <<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2012:194:0039:0043:IT:PDF>>
- Meaningful Bibliographic Metadata (M2B). Recommendations of a set of metadata Properties and encoding Vocabularies: <<http://aims.fao.org/fr/metadata/m2b>>
- Peter Suber's Open Access Policy options: <<http://legacy.earlham.edu/~peters/fos/newsletter/02-02-09.htm>>

- CRUI, *Open Access e la valutazione dei prodotti della ricerca scientifica*, 2009: <<http://www.crui.it/HomePage.aspx?ref=1782#>>
- from the EUA Working Group on Open Access adopted by the EUA Council, 2008: <http://www.eua.be/fileadmin/user_upload/files/Policy_Positions/Recommendations_Open_Access_adopted_by_the_EUA_Council_on_26th_of_March_2008_final.pdf>
- COMMUNIA Policy recommendation: <<http://communia-project.eu/final-report/>>
- EUROHORCs' on Open Access: <http://www.eurohorcs.org/SiteCollectionDocuments/EUROHORCs_Recommendations_OpenAccess_200805.pdf>

RANKING Web of Repositories: <<http://repositories.webometrics.info/>>

RAPPORTI:

- Automated downloading of citation data, COAR (Confederation of Open Access Repositories): <<http://www.coar-repositories.org/working-groups/repository-content/preliminary-report-sustainable-best-practices-for-populating-repositories/4-automated-downloading-of-citation-data/>>
- Best practice report on cultural heritage, Linked Heritage Project: <<http://www.linkedheritage.eu/getFile.php?id=229>>
- Digital Library Technology and Methodology Cookbook, DL.org Project Report, 2011: <<http://bscw.research-infrastructures.eu/pub/bscw.cgi/d222825/D3.4%20Digital%20Library%20Technology%20and%20Methodology%20Cookbook.pdf>>
- Europeana Libraries, Aggregating digital content from Europe's libraries. D5.1. Report on the alignment of library metadata with the European Data Model (EDM): <<http://www.europeana-libraries.eu/documents/868553/leade085-34ac-487f-82af-d5cd2545e619>>
- Functional Requirements for Subject Authority Data (FRSAD). A Conceptual Model, IFLA Working Group on the Functional Requirements for Subject Authority Records (FRSAR), 2010: <<http://www.ifla.org/files/assets/classification-and-indexing/functional-requirements-for-subject-authority-data/frsad-final-report.pdf>>
- Helping to Open up: improving knowledge, capability and confidence in making research data more Open, Report, Research Information and Digital Literacies Coalition (RIDLs), July 2013: <http://www.researchinfonet.org/infolit/ridls/strand5/>
- How to implement coordinated Open Access Policies in line with EC initiatives? MedOANet Report on the European Workshop Released, <http://libereurope.eu/news/how-to-implement-coordinated-oa-policies-in-line-with-ec-initiatives>
- IFLA Cataloguing Section Working Group on the Use of Metadata Schemas, IFLA (International Federation of Library Associations and Institutions): Final Report, 2005: <http://www.ifla.org/files/cataloguing/pubs/metadata_schemas-20050731.pdf>
- KeepIt Project. Setting institutional repositories on the path to digital preservation. Final project report from the JISC KeepIt Project, 2011: <<http://ie-repository.jisc.ac.uk/553/1/finalreport-keepit10.pdf>>
- Long-Lived Digital Data Collections: Enabling Research and Education in the 21st Century. United States. National Science Board: <http://www.nsf.gov/nsb/documents/2005/LLDDC_report.pdf>

- NISO Launches New Initiative to Develop Standard for Open Access Metadata and Indicators, 2013: <<http://www.niso.org/publications/newsline/2013/newslinefeb2013.html#report2>>
- Open Access - Fifth Report of Session 2013–14, House of Commons Business, Innovation and Skills Committee, UK: <<http://www.publications.parliament.uk/pa/cm201314/cmselect/cmbis/99/99.pdf>>
Open data e EU funding, Report, 2013:
- <<http://www.epsiplatform.eu/sites/default/files/150100928-2013-06-Public-Funding-and-Open-Data.pdf>>
- Report of the European Commission - Public Consultation on Open Research Data, October 2013: <http://ec.europa.eu/research/science-society/document_library/pdf_06/report_2013-07-open_research_data-consultation.pdf>
- Report of the European Commission - National Open Access and Preservation Policies in Europe, E.C., December 2011: <http://ec.europa.eu/research/science-society/document_library/pdf_06/open-access-report-2011_en.pdf>
- Research LibraryUK/SCONUL pubblicano un brevissimo report (curato da Alma Swan) dell'impatto dell'OA su università, ricercatori e società: <<http://www.rluk.ac.uk/content/rluksconul-briefing-impact-open-access>>
- Science-at-risk-NDIIPP-report: <<http://digitalpreservation.gov/meetings/documents/othermeetings/science-at-risk-NDIIPP-report-nov-2012.pdf>>
- Use of Content in Linked Heritage and Europeana”, LINKED HERITAGE Report, 2011: <<http://www.linkedheritage.eu/index.php?en/177/training-material-targeted-to-linked-heritage-content-providers>>

RDF/XML Syntax Specification (revised): <<http://www.w3.org/tr/REC-rdf-syntax/>>

RDF Vocabulary Description Language 1.0: RDF Schema: <<http://www.w3.org/tr/rdf-schema/>>

RDF & OPEN LINKED DATA un primo approccio – THGenius la ricerca nel Web Semantico. Blog “Find Information Share Knowledge”, 08-03-2011: <<http://atcult.wordpress.com/tag/linked-data/?blogsub=confirming#subscribe-blog>>

Registry of Open Access Repositories (ROAR): <<http://roar.eprints.org/>>

Reference Model for an OPEN ARCHIVAL INFORMATION SYSTEM - OAIS: <<http://public.ccsds.org/publications/archive/650x0m2.pdf>>

Regole italiane di catalogazione : REICAT / a c. della Commissione permanente per la revisione delle regole italiane di catalogazione. - Roma : ICCU, 2009: <http://www.iccu.sbn.it/upload/documenti/REICA_bozza_complessiva_genn2009.pdf>

Requisiti funzionali per i dati di Autorità. Un modello concettuale, a c.d. Glenn E. Patton, IFLA Working Group on Functional Requirements and Numbering of Authority Records (FRANAR). Edizione italiana a c. dell'Istituto Centrale per il Catalogo Unico delle biblioteche italiane e per le informazioni bibliografiche (ICCU), Roma, 2010: <http://www.ifla.org/files/assets/cataloguing/frad/frad_2009-it.pdf>

RESEARCH DATA, University of Bath: <<http://www.bath.ac.uk/research/data>>

(RESEARCH DATA is any material collected, observed, or created for the purposes of analysis to generate original research results, irrespective of the format of data. It may be digital, paper based or in other forms")

RESEARCH DATA MANAGEMENT nell'Università di Oxford,
<<http://www.bath.ac.uk/research/data/>>

Rete cooperativa di partner per il miglioramento dell'accesso mondiale all'informazione e alla conoscenza; dedicato in particolare alle opportunità per i paesi in via di sviluppo:
<<http://www.inasp.org.uk/index.html>>

RePEc (Research Papers in Economics): <<http://repec.org/>>

Repository disciplinare Econstor: <www.econstor.eu>

Repository ZENODO: <<http://www.zenodo.org>>

Repositories WEBOMETRICS: <<http://repositories.webometrics.info/>>

Research Excellence Framework (REF):
<<http://www.hefce.ac.uk/news/newsarchive/2013/name.78750,en.html>>

RESEARCH ITALY, il portale del MIUR per la ricerca italiana:
<<https://www.researchitaly.it/conoscere/stampa-e-media/news/verso-un-network-italiano-dell-open-access/>>

REST (Representational State Transfer). L'architettura Software architecture per sistemi distribuiti incluso servizi web RESTful:
<http://en.wikipedia.org/wiki/Representational_state_transfer>

Ricerca nel campo delle TIC: preservare il patrimonio digitale europeo a beneficio delle generazioni future"; EUROPA Press releases RAPID, 07/09/2010: <http://europa.eu/rapid/press-release_IP-10-1094_it.htm>

ROAR (Registry of Open Access Repositories): <<http://roar.eprints.org/>>

ROARMAP (Registry of Open Access Repository Material Archiving Policies):
<<http://roarmap.eprints.org/>>

Rome Declaration on CRIS and OAR:
<https://docs.google.com/document/d/1yvy8JzOQtoTij1Y8zjMlzUPWTBQsPJF-vlTixpJNQjA/edit?hl=en_US&pli=1; <http://cibernewsletter.caspur.it/?p=13895>>

Schema Bib Extend Community Group: <<http://www.w3.org/community/schemabibex/>>

SCHEMA.ORG: <<http://schema.org/>>; <<http://schema.org/docs/full.html>>

SCIDIP-ES - SCIENCE Data Infrastructure for PRESERVATION Project: <<http://www.scidip-es.eu/>>

Scimago Research Group. The Atlas of Science project: <<http://www.scimago.es/projects.php>>

Semantic Web Architecture: <<http://obitko.com/tutorials/ontologies-semantic-web/semantic-web-architecture.html>; http://en.wikipedia.org/wiki/Semantic_Web>

Semantic Web in Libraries Conference. SWIB13, 25 - 27 November, 2013, Hamburg, Germany: <<http://swib.org/swib13/programme.php>> (Introduction to Linked Open Data; Metadata Provenance Tutorial; Automatic Creation of Mappings between Classification Systems for Bibliographic Data; Cross-Lingual Semantic Mapping of Authority Files)

SHAMAN Project - Sustaining Heritage Access through Multivalent Archiving: <<http://shaman-ip.eu/>>

SKOS Simple Knowledge Organization System Primer: <<http://www.w3.org/TR/skos-primer>: <http://www.w3.org/TR/2009/REC-skos-reference-20090818>>

SKOS Use Cases and Requirements: <<http://www.w3.org/TR/2009/NOTE-skos-ucr-20090818/>>

SKOS wiki: <<http://www.w3.org/2001/sw/wiki/SKOS/Validation>>

Service Delivery Platform (SDP): <http://en.wikipedia.org/wiki/Service_delivery_platform>

SHERPA/RoMEO: <<http://www.sherpa.ac.uk/romeo>>

Sherpa_Politiche: <<http://www.sherpa.ac.uk/romeo/PDFandIR.php?la=en>>; <<http://www.sherpa.ac.uk/romeo/statistics.php?la=en&fIDnum=|&mode=simple>>

SHERPA JULIET: <<http://www.sherpa.ac.uk/juliet/>>

Semantica e Terminologia nei portali, Tavola rotonda, Firenze, 11-12 Dicembre 2000: <http://e-prints.unifi.it/archive/00000155/00/Semantica_e_terminologia_nei_portali.pdf>

Singapore Framework for Dublin Core Application Profiles, 2008: <<http://dublincore.org/documents/singapore-framework>>

Software TagTeam per l'aggregazione nelle piattaforme delle riviste editoriali tramite RSS: <<http://bit.ly/tagteam-intro>>

SPARC: <<http://www.arl.org/sparc/>>

SPARC, Selected list of Institutional Repositories: <<http://www.arl.org/sparc/repos/ir.html>>

SPARC Europe: <<http://www.sparceurope.org/>>

SPARQL Query Language for RDF: <<http://www.w3.org/TR/rdf-sparql-query/>>; <http://semanticweb.com/introduction-to-sparql_b22498; <http://www.w3.org/TR/2013/REC-sparql11-overview-20130321/>>; <<http://www.linkeddatatools.com/querying-semantic-data>>

Special issue on the state of the art in assessing research impact, in "Research Evaluation", Vol. 20, Issue 3, Sept. 2011. Guest editor Claire Donovan: <<http://rev.oxfordjournals.org/content/20/3.toc>>

Specification of metadata profiles and mappings to existing technology. VOA3R Agres AP Metadata Terms: <<http://www.ieru.org/voa3r/wiki/images/b/bb/20120913-D3.5-PartB.pdf>>

State of the Art and Roadmap for Current Research Information Systems and Repositories: <<http://www.irpps.cnr.it/it/eventi/workshop-on-cris-cerif-and-institutional-repositories#whitepaper>>

STATEMENT:

- sull'Accesso Aperto ai risultati della ricerca scientifica in Italia, 21.03.2013: <<http://air.unimi.it/cms/position%20statement.pdf>>; <http://wiki.openarchives.it/index.php/Position_statement_Open_Access>;
- on Open Access to Scholarly Literature and Research Documentation dell'International Federation of Library Associations (IFLA): <<http://www.ifla.org/publications/ifla-statement-on-open-access-to-scholarly-literature-and-research-documentation>>
- congiunto (tra LIBER, COAR, il Progetto OpenAIRE) sull'Open Data Pilot del programma Horizon 2020: <http://www.openaire.eu/en/about-openaire/publications-presentations/publications/doc_details/585-horizon2020opendatapilot20130703final>
- from the EUA Working Group on Open Access dell'European University Association: <http://www.eua.be/Libraries/Page_files/EUA_WG_open_access_1.sflb.ashx>
- from the University of Oregon Libraries: <http://library.uoregon.edu/diglib/irg/SB_Role.html>

SURplus: <<http://www.cilea.it/index.php?id=surplus>>

SWAP (Scholarly Works Application Profile):

<<http://dublincore.org/usage/reviews/2009/swap/>>; DCMI Usage Board Review of Scholarly Works Application Profile: <http://www.ukoln.ac.uk/repositories/digirep/index/SWAP>; <<http://www.ukoln.ac.uk/repositories/digirep/index/SWAP#Background>>

Sweo Community. Projects: LinkingOpenData:

<<http://www.w3.org/wiki/SweoIG/TaskForces/CommunityProjects/LinkingOpenData>>

SWIB 12. Semantic Web in Libraries. Towards an international LOD library ecosystem. 26 - 28 November 2012, Cologne, Germany: <<http://swib.org/swib12/>>, <[Slides and videos online](#)>

The Current State of Open Access Repository Interoperability. Confederazione di Repository Open Access (COAR), 2012: <<http://www.coar-repositories.org/files/COAR-Current-State-of-Open-Access-Repository-Interoperability-26-10-2012.pdf>>

The Dramatic Growth of Open Access, December 31, 2011:

<<http://poeticeconomics.blogspot.com/2011/12/happy-2012-open-access-movement.html>>

The European Library, LINKED OPEN DATA : <<http://www.theeuropeanlibrary.org/tel4/iod>>

The European Libraries - qualificazione metadati:

<<http://www.theeuropeanlibrary.org/tel4/login?targeturl=%3Flocale%3Dit>>

The Institutional Repository: Benefits and Challenges (American Library Association, 2013):
<http://www.ala.org/alcts/sites/ala.org.alcts/files/content/resources/papers/ir_ch05.pdf>

The Optimal and Inevitable for Researchers:
<<http://www.ecs.soton.ac.uk/~harnad/Tp/resolution.htm#2>>

TIMBUS Project - Digital Preservation for Timeless Business Processes and Services:
<<http://timbusproject.net/>>

Timeline Template - How to Create a Timeline in Excel:
<<http://www.vertex42.com/ExcelArticles/create-a-timeline.html>>

Tim's 5 star. Open Data plan with examples: <<http://5stardata>>
Topic Maps - Data Model. ISO/IEC JTC1/SC34 Information Technology -Document Description and Processing Languages: <<http://isotopicmaps.org/sam/sam-model/>>

Towards better Access to scientific information: Boosting the benefits of public investments in research, Commissione Europea, 2012: <<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2012:0401:FIN:IT:PDF>>

TRAC (*Trustworthy Repositories Audit & Certification: Criteria and Checklist*), CRL (The Center for Research Libraries), OCLC (Online Computer Library Center), v. 1.0, 2007:
<http://www.crl.edu/sites/default/files/attachments/pages/trac_0.pdf>

Translations of DCMI Documents: <<http://dublincore.org/resources/translations/>>

Trusted Digital Repositories: Attributes and Responsibilities. Rapporto del Research Library Group: <http://www.oclc.org/content/dam/research/activities/trustedrep/repositories.pdf>

Trusted Digital Repositories & Trusted Professionals. Conferenza internazionale, Firenze, 2012:
<<http://www.rinascimento-digitale.it/conference2012-culturalheritageonline-materials.phtml>>

Tutorial: Introduction to Linked Open Data (LOD): <<http://dcevents.dublincore.org/IntConf/dc-2013/paper/view/195>>

Understanding Metadata", National Information Standards Organization NISO Press, 2004:
<<http://www.niso.org/standards/resources/UnderstandingMetadata.pdf>>

U-GOV. Architettura, tecnologia e applicazioni. White Paper, luglio 2007,
<http://www.cineca.it/sites/default/files/WP-UGOV_1_0.pdf>:
<<http://www.cineca.it/page/sistema-u-gov>>

U-GOV Ricerca: da strumento di supporto alla valutazione di Ateneo a fonte dati per gli archivi aperti: <http://www.biblio.polito.it/eventi/aib_OAI/wp-content/interventi/fabrizioluglio.pps>

U-GOV, Tecnologia e Architettura: <<http://www.cineca.it/page/tecnologia-e-architettura>>

Unified Modelling Language: < <http://www.uml.org>: <http://www.omg.org/spec/UML/>>

Unified Modeling Language UML: riferimenti di base, introduzioni e linee guida:
<<http://www.analisi-disegno.com/uml/uml.htm>>

Unified Modeling Language Specifications (UML), ISO/IEC 19501: 2005:
<<http://www.omg.org/spec/UML/ISO/19501/PDF>>

UNESCO “Policy guidelines for the development and promotion of Open Access”:
<<http://www.unesco.org/new/en/communication-and-information/resources/publications-and-communication-materials/publications/full-list/policy-guidelines-for-the-development-and-promotion-of-open-access/>>

UNESCO Thesaurus, pubblicato come Linked Open Data: <<http://skos.um.es/unescothes>>

Use Case Authority Data Enrichment, World Wide Web Consortium (W3C),
User Guide/Creating Metadata , Dublin Core Metadata Initiative:
<http://www.w3.org/2005/Incubator/ld/wiki/Use_Case_Authority_Data_Enrichment>

Using usage statistics to encourage deposits, COAR (Confederation of Open Access
Repositories):
<<http://www.coar-repositories.org/working-groups/repository-content/preliminary-report-sustainable-best-practices-for-populating-repositories/4-automated-downloading-of-citation-data/>>

Utilizzo degli Archivi Istituzionali per la Valutazione della ricerca: esempi e opportunità
(Morando Maddalena). Blog del Politecnico di Torino “OA, Archivi Istituzionali e valutazione
della ricerca”: <http://www.biblio.polito.it/eventi/aib_OAI/interventi>

V4OA (Vocabularies for Open Access) Project - *towards a common approach to metadata for
research materials (embargo periods; the legal rights of organizations or authors; whether an
organisation or individual has paid for an item; to agree on metadata that asserts whether or not
an item is currently free to read online)*: <<http://v4oa.net/2012/12/13/progress-in-the-v4oa-project/>>; <<http://v4oa.net/2013/10/02/v4oa-public-consultation-paper-released/>>

Virtual International Authority File (VIAF): <<http://viaf.org/>>

VIVO Ontology. Enabling National Networking of Scientists:
<<http://vivoweb.org/files/SemanticModeling.pdf>>

VOA3R (Virtual Open Access Agriculture & Aquaculture) Open Access Repository:
<<http://voa3r.eu/>>

WF4Ever Project - Preservation of Scientific Workflows: < <http://www.wf4ever-project.org>>;

W3C RDB2RDF Incubator Group, A survey of current approaches for mapping of relational
databases to RDF, 2009:
<http://www.w3.org/2005/Incubator/rdb2rdf/RDB2RDF_SurveyReport.pdf>

What are Institutional Repositories?

EnablingOpenScholarship (EOS): <http://www.openscholarship.org/jcms/c_6162/repositories>

What is Data Management, Minnesota University:
<<https://www.lib.umn.edu/datamanagement/whatdata>>

What is Linked Data?, AI3 (Adaptive Information Innovation Infrastructure):
<<http://www.mkbergman.com/447/what-is-linked-data/>>

White Paper - U-GOV Ricerca. Il sistema per la gestione della Ricerca di Ateneo, Cineca, 2010:
<http://www.cineca.it/sites/default/files/Whitepaper_U-GOV_Ricerca_S3-1_0.pdf>

Wiki sull'Open Access in Italia, creato con lo scopo di raccogliere le principali informazioni, offre link ed esempi di buone pratiche relativi all'accesso aperto nel mondo ma soprattutto in Italia, destinati ad autori, bibliotecari ed amministratori:
<http://wiki.openarchives.it/index.php/Cos'è_l'_Open_Access>

Wiki.Openarchives.it:
<http://wiki.openarchives.it/index.php/Dati_e_cifre_sull%27Open_Access_in_Italia_-_2012#Progetti_Europei_sull'E2.80.99Open_Access_in_cui_.C3.A8_coinvolta_l'E2.80.99Italia>

Wiki.Openarchives.it: Open Access in Italia:
<http://wiki.openarchives.it/index.php/Dati_e_cifre_sull%27Open_Access_in_Italia_-_2012#PLEIADI>

Wiki.Openarchives.it: Applicazione delle Linee Guida:
<http://wiki.openarchives.it/index.php/Applicazione_delle_linee_guida#Universit.C3.A0_degli_studi_di_Modena-Reggio_Emilia>

World Bank's Open Access Policy: <<http://live.worldbank.org/bank-open-access-policy-development-liveblog>>

WorldCat: <<http://www.worldcat.org/>>

World- Cat Local: <<http://www.oclc.org/worldcatlocal/default.htm>>

XML (eXtensibleMarkupLanguage): <<http://www.w3.org/XML/>>

Zeitschriftendatenbank (ZDB): <<http://dispatch.opac.d-nb.de/DB=1.1/>>

ZBW Labs: <<http://zbw.eu/labs/en/about>>

5 ☆ Open Data: <<http://5stardata.info/>>

LISTA DELLE FIGURE

Figura 1. Una rappresentazione grafica dei collegamenti RDF tra le unità informative di un Catalogo e tra <i>dataset</i> nella <i>Linked Open Data Cloud</i>	3
Figura 2. La piattaforma di ricerca “ <i>OpenAGRIS</i> ” [V. sito] che impiega le potenzialità del Thesaurus AGROVOC in modalità <i>Linked Data</i>	4
Figura 3. La formalizzazione di unità informative e di loro collegamenti entro le piattaforme <i>TripleStores</i>	5
Figura 4. “ <i>The circle of the gifts</i> ” (regalo di scambio): il paradosso della proprietà intellettuale nella comunicazione dell’informazione scientifica	7
Figura 5. Il modello funzionale <i>Open Archive Initiative: Dara e Service Provider</i>	26
Figura 6. Il Portale unificato degli Archivi aperti nazionali olandesi della ricerca NARCIS	27
Figura 7. Il servizio “ <i>Ranking Web of Repositories</i> ”	29
Figura 8. Il numero dei mandati <i>Open Access</i> adottati da diverse Istituzioni a livello globale (2003-2013)	30
Figura 9. Gli Archivi Istituzionali: le strade per il deposito volontario e mandatario	37
Figura 10. I collegamenti tra diversi moduli del servizio Surplus a supporto della “ <i>governance</i> ” dei processi e contenuti di un’Istituzione	45
Figura 11. La “catena del valore” della “filiera della ricerca” supportata dal servizio “U-GOV- Ricerca”	47
Figura 12. L’architettura del sistema informatico integrato U-GOV	48
Figura 13. La comunicazione tra i moduli “UGOV - Catalogo della Ricerca” e l’Archivio Istituzionale del Politecnico di Torino	49
Figura 14. Le informazioni sul “Diritto d’Autore” sul sito dell’Archivio della Ricerca AIR dell’Università degli Studi di Milano	52
Figura 15. La piattaforma PHAIDRA: un sistema di gestione delle collezioni digitali con le funzioni di <i>archiviazione a lungo termine</i>	53
Figura 16. La filiera del ciclo di vita dei contenuti di un Archivio e i processi coinvolti	55
Figura 17. Archivio Istituzionale “ <i>EconStor</i> ” dell’Università Humboldt (Berlino) con la certificazione DINI	57
Figura 18. La funzione di navigazione “ <i>DDRAG and DROP</i> ” basata sui campi concettuali definiti dal Thesaurus AGROVOC	100
Figura 19. Le sezioni di ricerca “ <i>Map View</i> ” e “ <i>Tags Cloud</i> ” potenziate dalle voci del	100

Thesaurus AGROVOC

Figura 20. <i>Linked Open Data Cloud</i>	101
Figura 21. La sezione di ricerca e navigazione del Portale “ <i>The European Library</i> ” - “ <i>Authority search</i> ”	101
Figura 22. Alcune modalità di rappresentazione dei contenuti digitali nelle infrastrutture di ricerca oggetto d’indagine	104
Figura 23. La tecnologia informatica per la critica del testo “ <i>Pinakes Text</i> ” (<i>Strumento di filologia computazionale</i>)	105
Figura 24. Alcuni componenti del “Profilo Tecnico” di un Archivio digitale di ricerca	112
Figura 25. Piattaforma AIMS (<i>Agricultural Information Management Standards</i>)	113
Figura 26 (a). Analisi comparativa di 11 <i>Software</i> (commerciali e <i>Open</i>) eseguita entro il progetto “ <i>Repositories Support Project</i> ”	115
Figura 26 (b). Mappatura di alcuni principali aspetti in base a cui è stata eseguita l’analisi comparativa dei <i>Software</i>	116
Figura 27. Rappresentazione grafica dell’architettura dei Software DSpace e Fedora	120
Figura 28. Esempio di uno schema concettuale di metadati “ <i>item-level</i> ”	126
Figura 29. Schema di interazione tra <i>Data</i> e <i>Service Provider</i> OAI	128
Figura 30. Tecnologia e standard per gli Archivi digitali	132
Figura 31. Pagina online del <i>Virtual International Authority File</i> (VIAF)	135
Figura 32. Pagina online della Classificazione del MiUR - SSD (Settori Scientifico-Disciplinari)	137
Figura 33. Pagina online della Classificazione Decimale Dewey	138
Figura 34. Pagina online della Classificazione Universale Decimale	138
Figura 35. Pagina online della Classificazione di <i>Library of Congress</i>	139
Figura 36. Pagina online del VEST Registry dei Vocabolari controllati per dominio e tipologia	142
Figura 37. Pagina online del Servizio <i>Linked Open Vocabularies</i> (LOV)	146
Figura 38. Pagina online del Catalogo digitale “ <i>OPEN LIBRARY</i> ”	149
Figura 39. Processi di <i>Knowledge Discovery e Management</i> attivati dal tool “ <i>Concept Mapper</i> ”	151

Figura 40. La mappatura del “ <i>STW Thesaurus for Economics</i> ” con il “ <i>Thesaurus for the Social Sciences</i> ” per definire i termini per la ricerca nella <i>Cloud of search terms</i>	152
Figura 41. <i>Open Archival Information System (OAIS) - Reference model</i>	155
Figura 42. Rappresentazione dei manoscritti digitalizzati nell’Archivio Istituzionale dell’Università di Valentia	168
Figura 43. “ <i>Deposit of old Thesis</i> “ nell’Archivio Istituzionale <i>DSpace@Cambridge</i>	168
Figura 44. Rappresentazione dei contenuti di una Tesi digitalizzata nell’Archivio PHAIDRA (Università di Vienna)	169
Figura 45. Faccette a supporto delle Categorie Fondamentali per la ricerca e navigazione nel Portale DRIVER	172
Figura 46. Ricerca e navigazione per associazioni semantiche tramite Software <i>AquaBrowser Library</i> . Sito del <i>Museo Galileo</i>	173
Figura 47. Interfaccia di navigazione semantica “ <i>ONTOLOGY NAVIGATION</i> “	174
Figura 48. <i>ATLAS of SCIENCE</i> : navigazione semantica per campi del sapere nelle scienze	174
Figura 49. Navigazione semantica per <i>TIMELINE</i>	175
Figura 50. Navigazione semantica per <i>MAP INTERFACES</i>	176
Figura 51. Archivio VOA3R: ricerca e navigazione nelle pagine dei Profili di Autori depositanti	176
Figura 52. Archivio arXiv.org: consultazione di contenuti correlati gestiti dal Sistema editoriale <i>Public Knowledge Project</i>	177
Figura 53. Schema “ <i>Tim’s 5 star</i> ” di distribuzione di <i>Linked Open Data</i> a 5 stelle	190
Figura 54. Albero decisionale “ <i>LODE-BD</i> ” e tabella esplicativa a supporto dei passi della codifica dei valori del metadato “ <i>Subject</i> ”	205
Figura 55. Valori: “letterali” e “non letterali” per qualificare le proprietà di metadati descrittivi risorse informative	206
Figura 56. Passi decisionali per la scelta di Licenze a supporto di pubblicazione di diverse risorse	208
Figura 57. Modellazione e implementazione di metadati secondo le strategie <i>LODE-BD</i> “ <i>Design-time</i> ” e “ <i>Run-time</i> ”	237
Figura 58. Esempificazione dei grafi RDF registrati in <i>TripleStore</i>	238
Figura 59. Ricerca e recupero delle risorse nella piattaforma <i>AGRIS</i> tramite l’integrazione “ <i>on the fly</i> ” dei valori del <i>Thesaurus Linked Open Data</i> -	240

AGROVOC

Figura 60. Universo di Standard di metadati " <i>Seeing Standards</i> "	324
Figura 61. Architettura a livelli del Web Semantico	329
Figura 62. Esempio di componenti di uno statement di asserzioni RDF "Soggetto-Predicato-Oggetto"	331
Figura 63. L'esempio dei grafi formati dalle triple <i>RDF</i> alla luce della scheda di un catalogo	332
Figura 64. Le tecnologie del Web Semantico	333
Figura 65. Componenti della tecnologia <i>Linked Data</i>	335
Figura 66. Rappresentazione delle connessioni tra diversi <i>dataset</i> nel <i>Linked Open Data (LOD) Cloud Diagram</i>	336
Figura 67. <i>Library Linked Data (LLD)</i> : verso l'integrazione nello spazio globale dei dati interconnessi <i>Linked Open Data Cloud</i>	338
Figura 68. <i>Topic Maps</i> nei costrutti del Web Semantico	342
Figura 69. Livello astratto di diversi <i>Topic</i> , con Istanze e Associazioni. Occorrenze di <i>Topic</i> appoggiati sugli identificatori Web URI	343
Figura 70. Associazione tra <i>Topic</i> su due livelli	343
Figura 71. Dai modelli concettuali alle teorie logiche dell' <i>Unified Modeling Language</i>	344
Figura 72. Gerarchia delle Classi di <i>Topic Maps</i> rappresentata con il linguaggio UML	345
Figura 73. UML per la progettazione dell'oggetto "Risorsa" dell'Archivio di ricerca VOA3R	347
Figura 74. La modellazione dell'oggetto "Materiale" di un Archivio di ricerca	348
Figura 75. Lo strumento " <i>Poseidon for UML</i> " per la strutturazione grafica " <i>machine-readable</i> " dei modelli di dominio	349
Figura 76. Famiglia dei modelli concettuali FRBR (<i>Functional Requirements for Bibliographical Records</i>)	352
Figura 77. Modello DCAP (<i>Dublin Core Application Profile</i>) per la modellazione di profili applicativi di metadati a supporto di diverse tipologie di contenuti	354
Figura 78. Profilo Applicativo del Modello di dominio SWAP (<i>Scholarly Works</i>)	356

Application Profile)

Figura 79. Il modello DC-SWAP esteso	356
Figura 80. Le entità del modello CSMD per l'acquisizione delle informazioni di alto livello su studi scientifici	358
Figura 81. Iniziative della Confederazione di Repository <i>Open Access</i> per l'interoperabilità tra contenuti digitali di Repository	361
Figura 82. CERIF (<i>Common European Research Information Format</i>)	362
Figura 83. Ontologia VIVO: rappresentazione dei ricercatori e dell'intero contesto in cui essi lavorano	363
Figura 84. <i>CITO (Citation Typing Ontology)</i> : espressione delle citazioni correlate relative a un articolo di ricerca	364
Figura 85. Modello OBOE (<i>Extensible Observation Ontology</i>): descrizione di dati e processi scientifici empirici	365
Figura 86. Le unità concettuali di una Nano-pubblicazione	366
Figura 87. Possibili aggregazioni di risorse informative digitali tramite OAI-ORE (<i>Open Archives Initiative Object Reuse and Exchange</i>)	367
Figura 88. Modello OAI-ORE: dalle "Risorse aggregate" alla "Mappa delle Risorse"	368
Figura 89. "Enhanced Publications" accessibili via il Portale e l'aggregatore dei contenuti di ricerca NARCIS	368
Figura 90. <i>OpenAIRE's 'Enhanced Publications' Pilots</i>	369
Figura 91. Albero decisionale "LODE-BD" per la scelta delle decisioni sulla concettualizzazione del profilo del metadato SOGGETTO	372
Figura 92. Portale VOA3R: partner, contenuti, servizi e modelli sociali a supporto della diffusione dei contenuti di ricerca	377
Figura 93. Infrastruttura tecnologica del Repository di ricerca <i>Open Access</i> - VOA3R in vista del potenziamento dei canali comunicativi dei contenuti	378
Figura 94. Funzionalità di ricerca e navigazione dei contenuti nel <i>Repository</i> VOA3R	379
Figura 95. Sito dell'esperimento VOA3R " <i>Controlled Experiment for Reviewing Lifecycle Services</i> ": processo alternativo di revisione dei contenuti di ricerca in rete	380

LISTA DELLE TABELLE

Tabella 1. Caratteristiche e funzionalità principali di un Archivio aperto Istituzionale	35
Tabella 2. Punti principali da considerare in una Politica (<i>Policy</i>) dell'Archivio Istituzionale	36
Tabella 3. Alcune Raccomandazioni che possono essere applicate per la programmazione e la gestione degli Archivi Istituzionali	39
Tabella 4. Alcune Linee Guida a supporto di una migliore programmazione e gestione dei contenuti negli Archivi <i>Open Access</i>	41
Tabella 5. Scopi del sistema SURplus (Sistema Università Ricerca plus) <i>Open Archive</i>	44
Tabella 6. Funzionalità del sistema SURplus (Sistema Università Ricerca plus) <i>Open Archive</i>	45
Tabella 7. Passi d'indagine sulle piattaforme di ricerca dei contenuti digitali	60
Tabella 8 (a, b). Introduzione delle infrastrutture di ricerca scelte come oggetto di indagine	62 63
Tabella 9. Aspetti valutativi per indagare le piattaforme di <i>Repository</i> e Portali	66
Tabella 10. Risultati dell'indagine e alcune raccomandazioni per potenziare le piattaforme di ricerca dei contenuti digitali	91
Tabella 11. Maggiori fattori che possono catalizzare il valore e il successo di Archivi Istituzionali	107
Tabella 12. Schemi di metadati supportati da diversi <i>Software</i>	129
Tabella 13. Fasi o macroprocessi di filiera coinvolti nel Ciclo di vita di una risorsa in un Archivio digitale	178
Tabella 14. Alcuni Vocabolari controllati e Schemi semantici per la normalizzazione e l'abilitazione dei metadati d'Archivio come <i>Linked Data</i>	193
Tabella 15. Alcuni autorevoli documenti considerati dalle " <i>Linee Guida per la creazione e la gestione di metadati nei Repository Istituzionali</i> " (CRUI, 2012)	202
Tabella 16. Alcuni vocabolari controllati per l'abilitazione dei metadati " <i>RightsHolder</i> ", " <i>Licence</i> ", " <i>AccessRights</i> ", " <i>Rights</i> " come <i>Linked Data</i>	210
Tabella 17. Schema di metadati per la descrizione dei contenuti di ricerca negli Archivi Istituzionali basato sulle Linee Guida CRUI e le Raccomandazioni "LODE-BD"	216
Tabella 18. Tipologie di metadati in rapporto alle loro funzioni	320

Tabella 19. Classi o categorie concettuali di metadati <i>Dublin Core</i>	327
Tabella 20. Termini di metadati proposte dalle “Linee Guida per la creazione e la gestione di metadati nei <i>Repository</i> Istituzionali”	358
Tabella 21. Requisiti per la programmazione dei requisiti di metadati estensibili di qualità (il caso dell’Archivio VOA3R)	374

APPENDICI

Alcune modalità per la rappresentazione, gestione e l'interoperabilità dei contenuti digitali

Nelle APPENDICI (A, B) sono presentate alcune modalità per la rappresentazione, gestione e l'interoperabilità dei contenuti digitali, come:

1) Metadati, intesi come *dati sui dati*, *indici descrittivi*, *metatag* che rappresentano il contenuto, le proprietà e le funzionalità di una risorsa informativa. Metadati, intesi come informazioni comprensibili da una macchina e che possono essere estratti da una risorsa o trasferiti in rete con il documento (*record di metadati*) che accompagnano quest'ultima.

2) Metadati come *etichette descrittive* delle risorse informative aventi l'obiettivo di integrarle nello spazio del Web Semantico, un'estensione naturale del Web tradizionale che utilizza *schemi di metadati* o Ontologie per rappresentare domini di informazione attraverso le *classi* e *relazioni*, affinché la ricerca delle informazioni da parte degli utenti diventi più sofisticata, in quanto basata sull'interoperabilità di dati.

Gli *schemi di metadati* sono elaborati e distribuiti dai fornitori di informazione sulla base delle *caratteristiche specifiche* degli *oggetti* (e del loro ciclo di vita) da gestire e da utilizzare entro i sistemi di conservazione, di ricerca e di recupero di informazioni.

3) Tecnologie *Linked Data* (Web 3.0) e *Topic Maps* adoperate per la realizzazione del Web Semantico;

4) Potenzialità del linguaggio grafico UML (*Unified Modeling Language*) che permette - tramite l'utilizzo di modelli visuali - di analizzare, descrivere, specificare e documentare sistemi informativi di diversi domini;

5) Alcuni aspetti pratici del *Social Web* (Web 2.0) a favore di una maggiore cooperazione di contenuti digitali di varie comunità di utenti.

APPENDICE A

Metadati e alcuni modelli per la rappresentazione e la comunicazione dei contenuti digitali

I metadati costituiscono un corredo essenziale per la rappresentazione e la gestione delle risorse digitali, anche se essi “non debbono necessariamente essere percepiti come informazioni digitali” perché “il retaggio culturale e i professionisti dell'informazione da sempre hanno creato metadati, i quali fino ad oggi sono stati gestiti all'interno delle collezioni tradizionali” [De Robbio, 2001, online, non impaginato] come, ad esempio, nelle schede cartacei dei Cataloghi.

Nell'ambiente digitale i metadati sono comprensibili dalla macchina e costituiscono un tipo di informazione che può essere utilizzata in maniera appropriata dagli agenti Software.

Lo scenario del Web Semantico amplifica la concezione di (meta)dati posti per la ricerca dell'informazione in rete, spaziando oltre la loro interoperabilità di primo livello, quella sintattica (identificazione delle parole), e rinviandola a quella semantica (identificazione delle parole rispetto a classi, o concetti racchiusi negli schemi di diversi domini) importante per l'interoperabilità e l'ottimizzazione di organizzazione e ricerca dei contenuti. In tale contesto, i metadati si offrono come gli *accessi semantici* alla rete dell'organizzazione della conoscenza e, in particolare, ai documenti che la trasmettono tra diverse strutture di dati e ‘piani di espressioni’ che nell'ambito del Web Semantico vengono supportati da diverse tecnologie semantiche capaci di programmare e generare applicazioni e servizi ad alto valore aggiunto. La semantica dei dati consiste nel dare alla macchina delle informazioni utili in modo che essa possa utilizzare i dati nel modo corretto, convertendoli eventualmente.

that there can be no broader Definition of Semantics than the one we encounter in the Semantic Web, and furthermore, it has taken us significantly away from the original idea of Semantics - the linguistic definition of the term [...] if someone talks about providing a 'semantic solution' this should be viewed with caution, as it is open to widely differing interpretations, and is therefore somewhat meaningless without further qualification. [Stevenson, 2009, online, non impaginato]

1.1. Strategie di metadati nella ricognizione di aspetti strutturali e semantici di contenuti digitali

I metadati sono le informazioni (comuni e specifiche) che, associate ad un *oggetto informativo* o ad una collezione di essi, ne descrivono il *contenuto* e il *contesto* di riferimento e sono necessari per migliorare la loro *comunicazione* in termini di manifestazione, fruizione, recupero e conservazione (sia temporale che spaziale).

Secondo lo Standard¹ ANSI/NISO 3919-2005, il documento essendo un oggetto informativo viene definito come *Content Object*:

oggi il concetto di documento non fa, quindi, più esclusivo riferimento alla sua collocazione spazio-temporale quanto piuttosto ai *contenuti* in esso memorizzati, alla rappresentazione dei fatti o degli atti di cui esso è testimonianza. Ciò tanto più nei documenti digitali immateriali e non immediatamente intelligibili [...]

comunicazione ha trovato altre *forme* importanti di *rappresentazione* negli strumenti di *collaborazione* che progressivamente si integrano sulle *piattaforme web* [...] che in ragione di loro immediatezza, e dell'apparente veste informale sono in grado di superare quelle barriere alla *condivisione delle informazioni* esistenti [...] rivelando [...] loro integrazione in un *Sistema di gestione della Conoscenza*" [Rovella, 2008, pp.171, 181-182.]

Secondo lo Standard per la gestione documentale ISO 15489 [Micchetti, 2007], i *metadati* sono i dati che descrivono il *contesto*, il *contenuto* e la struttura dei documenti e la loro gestione nel tempo. E' necessario determinare *quali metadati* debbano essere creati insieme al documento e nel corso di processi documentali, e come *essi siano collegati e gestiti* in modo permanente.

La creazione di metadati può avvenire:

1) in modo automatico (es. estrazione di metadati dai file formattati in HTML/XHTML; generazione di metadati attraverso le tecniche di *clustering* [V. sito: *Insight Discoverer™ Clusterer*]);

2) semi-automatico, con la mediazione umana (tramite moduli elettronici appositamente progettati con le stringhe che devono essere compilate dai creatori o utilizzatori delle risorse) [V. sito: *Dublin Core Metadata Template*]);

3) manuale (un'unica concisa descrizione di *documenti/record* primari indicizzati attraverso l'analisi concettuale con le stringhe descrittive contenute nei documenti stessi.

valenza del termine *documento* che ha conosciuto e conosce una pluralità di accessioni e di significati che stanno, in gran parte, alla base delle diversificazioni esistenti. Anche all'interno delle scienze del libro e del documento non vi è una accettazione univoca della concettualità del termine che oscilla dall'accentuazione giuridico formale di parte di un *tutto* avente vita propria solo in quanto componente il *fascicolo* – unità primaria – a quella più estesa e più generale – mutuata da Otlet – di *oggetto informativo completamente indipendente dal supporto scritto* e quindi *insieme degli elementi descrittivi di una struttura testuale*. Sulle diverse concettualizzazioni del termine si giocano anche i rapporti – non sempre lineari – tra le varie discipline, rapporti spesso complicati dalla non condivisione di uno stesso vocabolario [Guarasci, 2008, p. 20].

I metadati generati possono essere:

1) aggiunti direttamente al documento;

2) associati a quest'ultimo essendo memorizzati in un *Repository* o database di metadati o in un file separato.

¹ *Standard*: documento riconosciuto con generale consenso o con l'approvazione di un organismo ufficiale.

Indipendentemente dal modo di creazione e dall'uso dei metadati, essi hanno la natura dinamica, cioè sono mobili nel tempo e nello spazio e continuano ad accrescersi entro il sistema entro cui dimorano, durante il ciclo di vita dell'oggetto informativo col il quale si relazionano.

È più che ragionevole cercare di condurre i metadati generati ad una forma normalizzata e standardizzata: i sistemi informativi, basati sui metadati standardizzati, migliorano le prestazioni di ricerca, forniscono l'individuazione più precisa della localizzazione e della disponibilità di un oggetto digitale e ne potenziano l'accesso, interoperabilità e gestione nel tempo [Chan, Zeng, 2006; Moffat, 2006].

1.1.1. Metadati: funzioni e tipologie

Diversi tipi di metadati (descrittivi, strutturali, amministrativi, sulla conservazione, tecnici, di utilizzo), ciascuno dei quali è utile per supportare una serie di funzioni assegnate, costituiscono il valore aggiunto delle risorse informative, orientate all'erogazione di contenuti e di servizi e non solo alla mera produzione di dati [V. sito: *Understanding Metadata*].

Osservando i metadati dal punto di vista delle funzioni che essi svolgono, si cita Greenberg [2006], secondo cui i metadati sono i *dati strutturati* dell'oggetto che supportano le *funzioni* ad esso associate. Nella Tabella 18 vengono riportate le tipologie di metadati rappresentate nel lavoro di *Greenberg* che riprende le ricerche sull'argomento fatte dagli autori come Lagoze, Gilliland-Swetland, Greenberg, Caplan. Tale rappresentazione sarà completata con la loro categorizzazione dei metadati secondo l'IFLA [V. sito: *IFLA Cataloguing Section Working Group on the Use of Metadata Schemas*].

Lagoze (1996)		Tipologie di metadati secondo alcuni autori e l'IFLA				
Funzionalità di metadati. Questi tipi di metadati facilitano:		Lagoze (1996)	Gilliland-Swetland (2000)	Greenberg (2001)	Caplan (2003)	IFLA (2005)
1.	Ritrovamento delle risorse. Ricerca informativa	descrittivi; identificativi	descrittivi	esplorativi	descrittivi	descrittivi
2.	Gestione delle risorse	amministrativi	amministrativi; di conservazione	amministrativi	amministrativi	amministrativi; altri metadati
3.	Utilizzo delle risorse	termini e condizioni d'uso	amministrativi; di conservazione; di utilizzo	di utilizzo tecnico; di uso intellettuale; amministrativi	amministrativi; di collegamento	di conservazione; di utilizzo; altri metadati
4.	Risorse in relazione all'utilizzo dall'utenza	di contenuto	di utilizzo	di utilizzo tecnico; di uso intellettuale	amministrativi; di collegamento	di utilizzo; di gestione dei diritti
5.	Autenticazione delle risorse, altre attività relative alla provenienza	di provenienza	amministrativi; di utilizzo	di autenticità; amministrativi	amministrativi	amministrativi; analitici
6.	Collegamento con altre risorse	di collegamento/relazionali	amministrativi	di autenticità; amministrativi	di collegamento	strutturali
7.	Esigenze di risorse hardware e software	strutturali	tecnici; di utilizzo	di utilizzo tecnico	strutturali	strutturali; tecnici

Tabella 18. Tipologie di metadati in rapporto alle loro funzioni

Nella Tabella 18 i *tipi di metadati* proposti da Gilliland-Swetland, Greenberg, Caplan e IFLA sono stati mappati con le funzionalità di metadati definite da Lagoze [1996], in corrispondenza con differenti connotazioni a seconda del ruolo al quale essi sono preposti o in relazione alle *funzioni* che essi svolgono.

Attualmente l'IFLA riconosce nove tipologie di metadati:

1) METADATI DESCRITTIVI: informazioni (*es.* Autore, Titolo; metadati *Dublin Core*) che descrivono le proprietà intellettuali, fisiche e il contenuto di un oggetto digitale, ne rendono possibile la presentazione, ricerca, l'individuazione e il recupero. Si trovano, generalmente, nelle basi dati dei sistemi di *Information Retrieval* all'esterno dell'Archivio digitale, e sono connessi a quest'ultimo tramite appositi collegamenti.

2) METADATI AMMINISTRATIVI: informazioni (*es.* data, acquisizione, diritti, reperibilità; metadati MAG [V. sito]) per facilitare la gestione amministrativa (l'organizzazione, la localizzazione, le statistiche) di risorse informative. Si tratta anche di informazioni fondamentali ai fini della conservazione e dell'accessibilità nel lungo periodo.

Per la conservazione dei *record* digitali i metadati amministrativi-gestionali si prestano particolarmente bene ad essere utilizzati all'interno di *modelli logico-funzionali* dell'Archivio degli oggetti digitali fortemente orientati alla conservazione permanente, come l'*Open Archival Information System* (OAIS) [Guarasci, 2010; V. sitografia: ISO 14721:2012].

Nel contesto di Biblioteche digitali, data la labilità dell'informazione, sono i metadati amministrativi e gestionali ad assumere un'importanza preponderante, rendendo infatti più affidabile e sicura l'informazione digitale, documentando le modalità di:

- *generazione e immissione* delle informazioni su oggetti digitali;
- *archiviazione e manutenzione* di questi ultimi nel sistema di gestione dell'archivio digitale.

In particolare lo schema di metadati MAG fornisce delle specifiche formali per le fasi di raccolta e archiviazione dei metadati e fornisce elementi per:

- *identificare in maniera univoca* gli oggetti digitali;
- *certificare l'autenticità* e l'integrità dei contenuti;
- *documentare la catena di custodia* degli oggetti digitali;
- *documentare i processi tecnici* eseguiti per la *conservazione permanente* degli oggetti digitali;
- *informare sulle condizioni e sui diritti di accesso* agli oggetti digitali da parte degli utenti finali.

I metadati amministrativi possono interagire con quelli descrittivi, ad esempio lo schema MAG importa il set di elementi *Dublin Core* (*Dublin Core Metadata Element Set/ DCMES* [V. sito]) tramite il *namespace* convenzionale *DC:* e li utilizza per la *descrizione* dell'oggetto *analogico* alla base della *digitalizzazione*; per esempio, nel caso della digitalizzazione di un volume cartaceo, gli elementi DC si occuperanno di registrare i metadati descrittivi relativi al medesimo volume. Inoltre, di recente si è cominciato a capire che un Archivio ben costruito rappresenta una fonte preziosa di dati amministrativi e gestionali che, opportunamente elaborati, possono dare una più *completa rappresentazione dell'attività di ricerca* svolta in un'Istituzione [Galimberti, 2010].

3) METADATI STRUTTURALI: informazioni che descrivono la tipologia e la struttura interna dei documenti (*es.* introduzione, capitoli, indice di un libro) e collegano le varie componenti degli oggetti digitali per un'adeguata e completa fruizione. Forniscono, quindi, i dati sui tipi, versioni, formati, connessioni, identificazione e localizzazione e altre caratteristiche di oggetti digitali e/o di loro componenti (*es.* di schemi di metadati SGML, XML, EAD, MOA2, METS ecc.)

4) METADATI ANALITICI: informazioni che analizzano e migliorano l'accesso al contenuto.

5) METADATI DI GESTIONE DEI DIRITTI: informazioni relative alle restrizioni (legali; finanziari, ecc.) in materia dell'accesso o dell'utilizzo di oggetti digitali.

6) METADATI TECNICI: informazioni sulle specifiche e sul funzionamento di particolari *Hardware* o *Software* usati per la digitalizzazione, la compressione, il recupero, il trasferimento e la rielaborazione di oggetti digitali.

7) METADATI DI CONSERVAZIONE: informazioni relative alla condizione fisica di risorse informative, alle azioni (*es.* *refreshing*, migrazione) volte alla loro preservazione e gestione nel tempo (*es.* metadati PREMIS).

8) METADATI DI UTILIZZO: informazioni relative al livello e al tipo d'utilizzo di elementi/oggetti/collezioni digitali.

9) ALTRI METADATI: particolari elementi di metadati basati sui requisiti locali, regionali, e/o organizzativi, o sugli standard nazionali di metadati.

Secondo le necessità della descrizione, strutturazione o gestione di diverse tipologie delle risorse informative, si potrebbero scegliere *metadati diversi*, adottare *prospettive diverse* e proporre *diverse distinzioni* e Classificazioni [Vanni, Fini., 2004].

L'esistenza di una *molteplicità tipologica di metadati* sarà il nodo cruciale per una *continuità di accesso e di utilizzo*, fisico e intellettuale, delle risorse digitali informative e degli oggetti informativi che esse contengono.

Come abbiamo notato, i fattori che connotano la *specificità dei metadati* riguardano il loro inscindibile *legame con gli oggetti informativi*, di cui essi veicolano le informazioni e ne consentono l'*indicizzazione* da parte di *strumenti di ricerca* appositamente predisposti per interpretarli.

L'*universo di metadati* si posa sulle *informazioni di contenuto* manifestato attraverso le risorse informative che, a loro volta, entrano a far parte delle *collezioni contraddistinte* a seconda dei diversi domini e di comunità di riferimento. La necessità di ricognizione strutturale e concettuale delle risorse informative nei diversi domini implica l'uso di metadati, o meglio degli Schemi (modelli) di metadati secondo i loro specifici obiettivi e connotazioni.

Il termine *schema*, generalmente inteso come struttura (*structured framework*) [V. sito: *Miriam-Webster online*], può essere visto come un *vincolo formale sulla sintassi*, per una rappresentazione consistente della *semantica dei dati*.

Di solito, gli Schemi di metadati vengono formalizzati da una certa autorità di registrazione, che definisce il *vocabolario di un particolare insieme di dati*, cioè nomi di elementi e regole di presentazione delle loro *proprietà*, delle *relazioni* con altri elementi, così che essi possano essere modellati e letti in modo uniforme da diversi utenti (persone e applicazioni).

Secondo Greenberg [2005], per *Schema di metadati* vengono considerati:

1. una *collezione di elementi di metadati* raccolti a sostegno di un oggetto informativo, supportando la sua *funzione* o una serie di funzioni (es. per la scoperta, l'amministrazione, la strutturazione, l'uso dell'oggetto informativo);

2. una collezione di *elementi di metadati* che formano un *contenitore strutturato*, in cui vengono aggiunti i *valori di dati*. Tali valori possono essere non controllati o controllati (ad esempio, tratti da una sorgente come *Library of Congress Subject Headings (LCSH)* o da un elenco di valori standardizzati);

3. una *collezione di elementi di dati* (o un vocabolario di dati) con i loro *attributi formalizzati in una specificazione* (nome, identificatore, etichetta, definizione, data ecc.).

In un *record di metadati* possono *coesistere metadati* provenienti da *schemi e tipologie di metadati diversi*. L'uso di diversi *schemi di metadati* deve essere coerente con le esigenze:

- di diverse comunità che li adottano;
- di specifiche tipologie e granularità della descrizione di risorse informative;
- dei *profili applicativi* per la descrizione delle risorse;
- e finalità dei Progetti per sviluppo e gestione delle collezioni.

[V. sito: *Digital Libraries: Metadata Resources*].

Le varie comunità di utenti definiscono e usano diversi *schemi di metadati* finalizzati a rappresentare nel modo più efficace gli oggetti informativi appartenenti ai domini di loro interesse e a soddisfare una grande varietà di requisiti funzionali di diversi sistemi informativi.

In alcuni casi gli *Schemi di metadati* possono essere definiti in modo *autonomo* (senza tener conto degli schemi già esistenti relativi a risorse delle medesima natura). Dall'altra parte, gli schemi esistenti potrebbero essere *ampliati* fino a *comprendere tutti i dati necessari* al soddisfacimento dei requisiti funzionali:

nessuno schema di metadati può, da solo, essere *valido* per qualsiasi applicazione, per tutti i generi di risorse e per ogni comunità di utenti. Le informazioni che lo descrivono debbono, quindi, essere ospitate in una *struttura flessibile e modulare*, in cui vengano registrati, nelle varie procedure, gli elementi ricavati, di preferenza, direttamente dal documento [Weston, 2007, slide].

Profili applicativi di metadati nascono dalle esigenze di avere un accesso integrato a risorse diverse, combinando insieme elementi di *diversi schemi di metadati*. Grazie alla definizione e documentazione di un *profilo di applicazione*, infatti, è possibile conoscere e sfruttare adeguatamente (ottimizzare) diversi elementi di una varietà egli schemi in un'applicazione e accedere agli oggetti digitali in modo preciso, integrando tra loro risorse e contribuendo ad una migliore comunicazione tra applicazioni diverse [V. sito: I profili di applicazione].

1.1.2. Universo di Standard di metadati

La maggior parte degli Schemi di metadati è stata riconosciuta dalle autorità di registrazione come gli Enti di standardizzazione. Un Elenco Ufficiale di tutti gli Schemi e Standard di metadati disponibili ancora non esiste, anche se il numero di Registri² di metadati sta diventando sempre più grande. Tra i più importanti registri di metadati si potrebbe citare:

- *UKOLN's schemas Registry*;
- *CORES Registry*;
- *JISC IE Metadata Schema Registry (IEMSR)*;
- *The European Library (TEL) metadata registry*;
- *DCMI Metadata Registry*;
- *DESIRE, SCHEMAS, Metaform*;
- *ULIS Open Metadata Registry*;
- *IFLA Digital Libraries: Metadata Resources web page*;
- *Global Information Locator Service (GILS)*.

Tali Registri, - rappresentando strumenti essenziali per documentare (in modo analitico, standardizzato e continuativo) diversi Schemi di metadati (riconosciuti anche come Standard), Norme della descrizione bibliografica, Vocabolari Controllati utilizzati per supportare il controllo di autorità [Guerrini, Tillett, 2003] dei valori di metadati.

Il numero degli Standard e Schemi di metadati nei diversi settori è travolgente, e il loro ritrovamento e le interrelazioni complicano ulteriormente la situazione della loro selezione e dell'uso appropriato nei specifici domini applicativi.

One key challenge for e-Government programs around the world has been the *lack of easily accessible information about the metadata schemas, controlled vocabularies, code lists, and other reference data* that provide interoperability among a broad diversity of data sources. The *Asset Description Metadata Schema* was developed for exchanging information about such "interoperability assets". The schema was developed with support from the European Commission with the objective of facilitating interoperability across eGovernment programmes in Europe, but it is already proving its usefulness in a wider context, for example to describe specifications maintained by DCMI and W3C [V. sito: Metadata for Public Sector Administration, Asset Description Metadata Schema].

Intanto, seppur diversi Schemi di metadati presentano molte differenze (per livello di specificità, struttura, maturità, tipo di utenza e di utilizzo), esiste anche affinità fra i metadati data dal fatto che il loro scopo principale è quello di descrivere, identificare, accedere e navigare i contenuti di risorse informative.

La piattaforma *MetaMap* [V. sito] ha lo scopo di aiutare diverse comunità delle scienze dell'informazione a comprendere l'esistenza di diversi Standard di metadati e iniziative pratiche in questo ambito. Il modello di visualizzazione *Seeing Standards: A Visualization of the Metadata Universe* (Figura 60) raggruppa in modo schematico gli Standard di metadati secondo le comunità, domini, funzioni e scopo, fornendo link per consultare il Glossario di alcuni importanti Standard di metadati [V. sito: *Glossary of Metadata Standards*].

² *Registri*: sistemi che registrano informazioni sull'esistenza, sulla creazione, sulla struttura, sulla semantica, sulla gestione, sul mantenimento di diversi schemi di metadati (possibilmente anche sui formati di scambio di ogni tipo di dato) ecc.

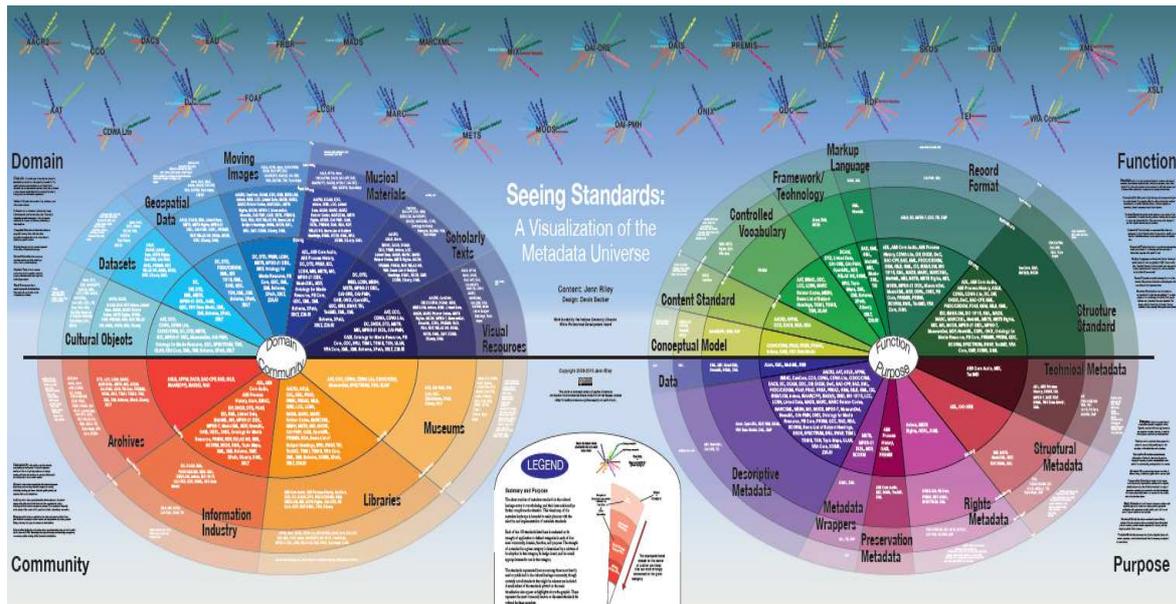


Figura 60. Universo di Standard di metadati "Seeing Standards"

Fonte: *A Visualization of the Metadata Universe*,

<http://www.dlib.indiana.edu/~jenlrile/metadatamap/>

Ciascuno dei 105 Standard presentati in Figura 60 viene valutato secondo la sua potenzialità di applicazione alle categorie definite su quattro assi:

- la comunità,
- il dominio,
- la funzione e
- lo scopo.

Tale potenzialità è determinata dal grado della adozione dello Standard in una certa categoria, dal suo intento di design del dominio, e dalla sua adeguatezza complessiva per l'uso in quella categoria. Norme, Modelli, Vocabolari, Schemi e Standard (Figura 60) a supporto della presentazione delle risorse informative sono tra quelli più utilizzati, comunemente riconosciuti e/o discussi nella comunità dei Beni Culturali:

- DC (*Dublin Core Metadata Element Set*);
- CCO (*Cataloging Cultural Objects Metadata Structure Standard*);
- EAD (*Encoded Archival Description Metadata Structure Standard*);
- VRA (*Visual Recourse Association Metadata Structure Standard*);
- CIDOC/CRM (*CIDOC Conceptual Reference Model Metadata Element Set*);
- MODS (*Metadata Object Description Schema*);
- MARC (*Machine Readable Cataloging Metadata*);
- GEM (*Gateway to Educational Materials Metadata*);
- TEI (*Text Encoding Initiative Metadata*);
- METS (*Metadata Encoding and Transmission Standard*);
- EAC-CPF (*Encoded Archival Context - Corporate Bodies, Persons, and Families*);
- FRBR (*Functional Requirements for Bibliographic Records*);
- LCC (*Library of Congress Classification*);
- BISAC (*Book Industry Standards and Communications Subject Headings List*);
- AACR2 (*Anglo-American Cataloging Rules*);

- RDA (*Resource Description and Access*);
- ISO/IEC 20943 (*Procedures for Achieving Metadata Registry Content Consistency*);
- ISO/IEC 20944 (*Metadata Registry Interoperability and Bindings*);
- METS (*Metadata Encoding and Transmission Standard*);

Tra gli altri Standard che assumono la responsabilità di certificare i contenuti informativi, di rappresentarli per la ricerca, di conservarli e di garantirne la preservazione a lungo termine, si possono anche citare anche:

- IEEE LOM (*Standard for Learning Object Metadata*);
- PRISM (*Publishing Requirements for Industry Standard*; definisce un insieme di elementi di metadati per esigenze comuni di pubblicazioni editoriali);
- MPEG-21 (*Moving Picture Experts Group*);
- MEI (*Music Encoding Initiative*);
- SMIL (*Synchronized Multimedia Integration Language*);
- ODRL (*Open Digital Rights Language*);
- OLAC (*Open Language Archives Community*) basato sul *Dublin Core* qualificato;
- PREMIS *Preservation Metadata* (metadati per la conservazione a lungo termine) e

altri ancora, ognuno di esse propone soluzioni specifiche per diverse esigenze di rappresentazione e supporto dell'informazione digitale.

La funzione normalizzatrice dei contenuti digitali svolta dai metadati è inoltre condizione per l'industrializzazione dei processi di digitalizzazione, e contribuisce a creare un mercato di prodotti e servizi in questo settore.

Negli ultimi anni l'ambiente digitale ha visto un'enorme crescita di *interazione, collaborazione* e scambio delle informazioni tra *diverse comunità di metadati* che in precedenza erano relativamente isolate, cioè costruite sulle necessità locali (istituzionali e quelli di utenza). Tuttavia, diversi Standard di metadati non sono stati ancora in grado di soddisfare le esigenze di interoperabilità tra le comunità di standardizzazione indipendenti.

Il concetto di *armonizzazione* di diversi Standard di metadati, definito come *l'interoperabilità delle combinazioni di diverse specifiche di metadati*, è emersa come una questione centrale per il *futuro di metadati basati sul web*, che rappresenta una sfida cruciale nella progettazione di *crosswalks (metadata mapping)* [V. sito: *Metadata Mapping between metadata formats*] per una fertilizzazione incrociata di diversi sistemi di metadati, affinché essi possano provvedere ad un efficace ed efficiente interscambio e riuso (sintattico, strutturale e semantico) dei metadati e affinché gli utenti siano più soddisfatti nel recupero delle informazioni attraverso una ricerca simultanea nei sistemi di metadati eterogenei [Chan, Zeng, 2006; Nilsson, 2010].

1.1.3. *Dublin Core Metadata Element Set*

Attualmente, uno tra i modelli (sistemi) di metadati più diffusi e condivisi a livello internazionale è il set di metadati strutturati *Dublin Core*, DC (*Dublin Core Metadata Element Set*, DCMES) [V. sito] che è il risultato di un convegno svoltosi nel 1996 a Dublin (Ohio), sede della rete OCLC (*Online Computer Library Center*) [V. sito], al quale i diversi partecipanti – Bibliotecari, Archivist, Editori, Ricercatori e sviluppatori di Software - hanno collaborato al fine di fornire strumenti per l'accesso alle risorse digitali appartenenti ai diversi domini applicativi [Guerrini, Ventura, 2007].

Quindi, il DC è il risultato di un consenso ampio ed *interdisciplinare sulla natura dei dati* indispensabili per la descrizione di *diversi oggetti digitali* distribuite in rete e per agevolare l'effettuazione della ricerca su essi. ed è a supporto alle descrizioni di diverso tipo di risorse.

Il DC è formalizzato con lo standard ISO 15836:2009 ed è adoperato in molti sistemi come *l'elemento semantico di base per altri tipi di metadati*, perseguendo l'obiettivo di interoperabilità e trasversalità ai fini di *favorire un approccio integrato alle informazioni*. Il continuo

aggiornamento del modello DC e l'incessante attenzione alla sua "qualità, il confronto sia con il mondo dell'università (biblioteconomia, informatica, ecc.), sia con quello della professione (Biblioteche, Archivi, Musei, ecc.) sono garantiti dalla *Dublin Core Metadata Initiative* (DCMI), un'organizzazione non profit con sede a Singapore" [Guerrini, 2010, p.36] che ha lo scopo di sviluppare *Standard interoperabili* per i contenuti online.

Il formato di metadati DC si caratterizza come uno strumento:

1) valido per la descrizione di *risorse informative di diverse tipologie*, persino di quelle che per la loro descrizione necessitano dei metadati specifici.

Per questo scopo si offre molto bene il modello astratto *Dublin Core Abstract Model* (DCAM) [V. sito] su cui per diversi domini di informazione è stata sviluppata una varietà di *profili applicativi* di metadati adeguati alle specifiche esigenze della descrizione.

A titolo di esempio si veda il Profilo applicativo SWAP (*Scholarly Works Application Profile*) [V. sito] per modellare la descrizione dei contenuti di ricerca scientifica. SWAP è basato sul *Dublin Core Abstract Model* e sulle specifiche FRBR (*Functional Requirements for Bibliographic Records*) [V. sito].

Il DC si sta proponendo, inoltre, come uno strumento collaborativo per descrivere il materiale catalografico nel mondo delle Biblioteche Digitali [Saffady, 1995; V. sito: *DC-Library Application Profile*] trovando, ad esempio, le aree di interazione con i metadati catalografici nel formato MARC [Gorman, 2004];

2) base per l'*estensione* della *descrizione semantica* dei valori dei propri metadati, affiancandoli (ove necessario) da *qualificatori* che possono derivare da Vocabolari Controllati e descrizioni Standard [V. sito: *Dublin Core Qualifiers*];

3) utile per la *descrizione* di oggetti digitali in *diverse lingue* (la semantica del DC è stata tradotta in varie lingue) [V. sito: *Translations of DCMI Documents*];

4) flessibile per sviluppare una *comune rete di dati strutturati concordati nel significato e valore*, appoggiandosi sul modello di dati RDF (*Resource Description Framework*) [V. sito], lo strumento base proposto da W3C (*World Wide Web Consortium*) per la strutturazione, codifica, lo scambio e il riutilizzo di metadati. Impiegando RDF nelle descrizioni delle proprie risorse, diverse applicazioni Web diventano in grado di interagire (comunicare) tra di loro, scambiano reciprocamente i contenuti che gestiscono;

5) adatto per contribuire – tramite l'elaborazione delle equivalenze tra i descrittori e le opportune relazioni tra elementi di metadati – ad una migliore interoperabilità tra contenuti di varie comunità come Archivi e Biblioteche digitali, mondo Editoriale, Musei, agenzie governative e così via.

Il DC viene ampiamente utilizzato dalle risorse che richiedono descrizioni complesse (e in diverse lingue), ponendosi anche come una sorta di *linguaggio descrittivo intermedio* che interviene nella *mappatura tra diversi modelli di metadati* e, di conseguenza, nei raffinamenti dei livelli della descrizione e dei meccanismi della ricerca.

L'adozione del set di metadati DC da parte di diversi gruppi di operatori socio-culturali:

- consentirà una più *puntuale identificazione* dei *contenuti digitali eterogenei* ma *relazionati* nei settori trasversali;
- faciliterà *convergenza, comunicazione e interoperabilità* (sintattica, semantica) [Trigari, 2007; Weston, 2007] tra vari applicazioni e sistemi e
- permetterà un uso professionale del DC nell'ambito degli strumenti di mediazione propri delle diverse discipline (*interoperabilità multidisciplinare*).

Nello specifico, il *Dublin Core* (DC) è un *sistema/modello di metadati* costituito da un nucleo di elementi essenziali (15 elementi di *Dublin Core Metadata Element Set* (DCMES) ai fini della *descrizione di qualsiasi materiale digitale* accessibile via rete informatica. DCMES permette la creazione di *descrizioni semplici e flessibili* (i 15 elementi sono opzionali e ripetibili), facilmente comprensibili e condivise da diverse applicazioni e comunità dell'informazione. Nonostante vi

sia la possibilità di essere esteso e personalizzato anche a sottoelementi o *qualificatori* (essendo abbastanza generale da includere ogni indicazione di contenuto semantico) per soddisfare le esigenze particolari, il DC mantiene una struttura stabile.

Secondo Taylor [2003], gli elementi DC sono suddivisi in tre gruppi (Tabella 19):

- I. Contenuto della Risorsa;
- II. Proprietà intellettuali della Risorsa;
- III. Caratteri peculiari.

I gruppo:		II gruppo:		III gruppo:	
Contenuto della Risorsa		Proprietà intellettuale della Risorsa		Caratteri peculiari	
Elementi di Dublin Core impiegati					
Titolo (Title)	Copertura (Coverage)	Creatore (Creator)	Data (Date)		
Soggetto (Subject)	Lingua (Language)	Editore (Publisher)	Tipo (Type)		
Descrizione (Description)	Relazione (Relation)	Autore di contributo subordinato (Contributor)	Identificatore (Identifier)		
Fonte (Source)		Gestione dei diritti (Rights)	Formato (Format)		

Tabella 19. Classi o categorie concettuali di metadati *Dublin Core*

Il DC risulta un formato di metadati che – oltre alla funzione descrittiva - ha supplito anche ad *altre funzioni* riconducibili sotto differenti, più specifiche tipologie di metadati.

Quanto al meccanismo di qualificazione, esso si basa sull'uso degli Elementi di Raffinamento (*Refinement Elements*) o Qualificatori (*Qualifiers*) che possono essere affiancati da diversi Schemi di Codifica (*Encoding Scheme*), Standard, Schemi di Classificazione e Vocabolari controllati in uso da diverse comunità di utenti.

Di seguito vengono presentati alcuni esempi di elementi DC affiancati dai *qualificatori* usati per riferirli e per restringerne l'ambito semantico:

```
date.ISO-8601=2011-12-05 T11.53CET;
date.created=August 2011; date.valid=September 2011; date.available=October 2011;
date.issued=November 2011; date.modified= Dicembre 2011;
dc:subject xml:lang="fr"; dc.subject.lcsh=Insects-Fossil-Lebanon.
```

Gli utilizzatori possono anche definire e applicare *qualificatori di uso locale*, purché rispettino la struttura generale, e *diffondere le varianti* locali che potranno diventare parte integrante dello Standard DC.

Vista la semplicità dell'impiego dei descrittori DC e il costante sostegno fornito dal sito web della comunità DCMI che offre informazioni e suggerimenti per l'applicazione dello Standard (la traduzione italiana delle specifiche descrittive dello standard DC e curata dall'ICCU [V. sito]), il DC può essere facilmente utilizzato da chiunque (sia da non catalogatori che da specialisti) ritenga la sua struttura adeguata a descrivere i propri documenti/collezioni di essi, anche se appartengono a varie discipline diverse (perseguito così gli scopi di interdisciplinarietà e di interoperabilità semantica) e sono scritti in diverse lingue.

Nelle piattaforme digitali di comunicazione dell'informazione scientifica supportate dal movimento *Open Access* e appoggiate sull'architettura *Open Archive Initiative* [V. sito] il DC viene selezionato come '*Core Metadata Format*' per permettere l'interoperabilità di base e la disseminazione dei contenuti scientifici entro l'architettura che connette diversi *Data Provider*

(*Repository*/Archivi di metadati e contenuti) e *Service Provider* (servizi centralizzati di *harvester* di metadati, aggregatori) che si occupano di scaricare i dati da un Archivio all'altro, per rendersi visibili ai fini della raccolta, esponendo i propri metadati al cosiddetto "colloquio" entro la comunità internazionale.

Il protocollo di interoperabilità *Open Archive Initiative Protocol for Metadata Harvesting* (OAI-PMH) – che può supportare insiemi paralleli di metadati, disseminati dagli Archivi distribuiti - facilita la localizzazione, la disseminazione e la raccolta dei metadati esposti da diversi sistemi di aggregazione dei dati e contenuti sul web. Il protocollo OAI-PMH prevede che i *Data Providers* con i loro Archivi di contenuti possono ritenere di definire *link* specifici all'interno del *metadato* verso il *contenuto*. In tale ottica il formato DC fornisce gli elementi identificatori che possono essere usati a questo scopo, essendo raccomandato come componente di base per lo scambio di informazioni e per supportare la ricerca di esse in domini applicativi diversi, contribuendo così in modo significativo al concetto di *interoperabilità*, la capacità di un sistema o prodotto di operare con altri sistemi senza che questo richieda sforzi particolari da parte dell'utente (*"the ability of a system or a product to work with other systems or products without special effort on the part of the custode"* <whatis.com>).

1.2. Web Semantico: nuovi approcci di comunicazione dei contenuti in rete

L'espressione "*Semantic Web*" è stata coniata nel 2001 da Tim Berners Lee (il creatore del *World Wide Web*) per definire l'estensione del Web tradizionale, attribuendogli la capacità di *comprendere* la *semantica* dei documenti digitali in rete e, in fase di ricerca, di interagire con l'utente nelle operazioni sofisticate di scoperta ed estrazione dei contenuti, nonché di elaborare nuove tecniche e nuovi *set* di basi di conoscenza [Zareen, Finin, 2011].

This Web is not a separate web but an extension of the current web in which the information is given well-defined meaning, better enabling computers and humans to work in cooperation. The prerequisite of this web, as its definition implies, is Metadata that explicitly represent Semantics of data, which called Ontology [Safari, 2004, p.7].

L'idea del progetto "Web Semantico" nasce semplicemente dall'ampliamento dell'idea di utilizzare *schemi di metadati* per descrivere interi domini di informazione, mappandone i dati rispetto a Classi o concetti.

Un punto di forza principale del Web Semantico è sempre stata l'espressione, sul Web, della grande quantità di informazioni del *database relazionale* formulate in una modalità *processabile da una macchina*. Il formato di serializzazione RDF [...] è un *formato funzionale* ad esprimere le *informazioni di database relazionale* [Guerrini, Possemato, 2012, p. 10]

tramite le triple (Soggetto/*Risorsa* – Predicato/*Proprietà*- Oggetto/*Valore*).

Il Web Semantico - tramite la sua architettura a livelli (Figura 61) - ha l'obiettivo di:

- *estendere la creazione di metadati* per sistemi di ricerca dell'informazione guidati dalla *semantica* basata sulle relazioni tra gli oggetti informativi digitali;
- amplificare l'interazione multifunzionale e anche multi-linguistica tra "uomo-macchina" e "macchina-macchina".

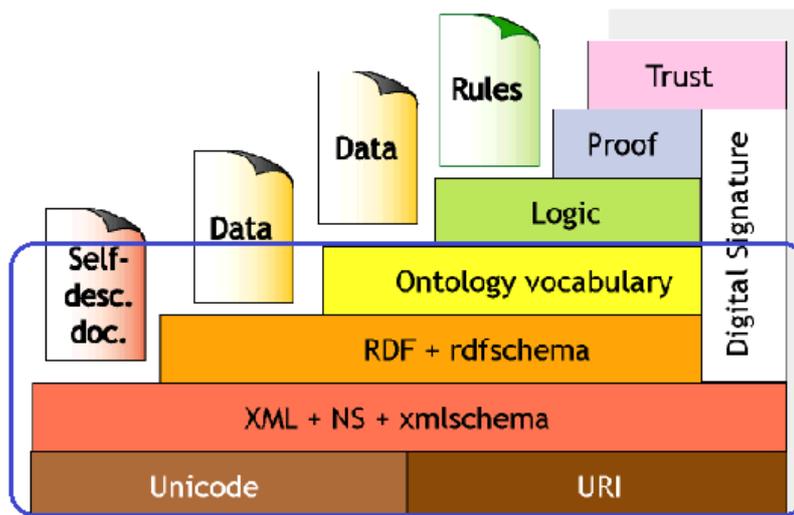


Figura 61. Architettura a livelli del Web Semantico

Le fondamenta architetturali del Web Semantico - dichiarate come tali dal *World Wide Web Consortium* (W3C) – sono seguenti:

- 1) lo standard URI (*Uniform Resource Identifiers*) per la definizione univoca di indirizzi Internet dei contenuti digitali;
- 2) il linguaggio XML (*eXtensible Markup Language*) che gioca un ruolo di base con i *namespace*³ e gli *XML Schema*. Con XML è possibile modellare in digitale la realtà che si considera;
- 3) sul RDF (*Resource Description Framework*), lo strumento base per la codifica, lo scambio e il riutilizzo di metadati strutturati sul Web per mezzo di:
 - *RDF Data Model* (con la sua codifica in XML) che esprime la struttura per l'espressione dei dati;
 - *RDF Schema* che permette di definire specifici vocabolari per i metadati e creare nessi tra oggetti

da disporre così di strutture in grado di descrivere e automatizzare i collegamenti esistenti fra diversi *dataset* sul Web;

- 4) sul livello Ontologico. Una Ontologia Web esprime le *relazioni* fra Concetti, che diventano *Classi di dati* coinvolte nella strutturazione della conoscenza di una certa realtà (processi, domini di informazione) [Gnoli, Scognamiglio, 2008]. Il genere più tipico di Ontologia per il Web è rappresentato da una Tassonomia ovvero uno strumento cognitivo utilizzato per classificare e categorizzare un insieme di cose tramite la definizione di Classi di oggetti e di relazioni tra loro, e da una serie di regole di inferenza.

Le singole applicazioni del Web Semantico possono operare processi di inferenza utilizzando *specifici vocabolari* definiti con i linguaggi formali e semi-formali comprensibili dai computer, interrogando così la semantica dei dati pubblicati sul Web.

Di conseguenza l'informazione deve in qualche modo diventare *machine-understandable*, assumere un livello di significato, o perlomeno di significante estremamente strutturato, anche per Software avanzati, nell'articolo denominati "Agenti", che siano in grado di effettuare *inferenze*, ottenendo così in maniera automatica nuove informazioni che verranno inviate e utilizzate dall'utente finale, o

³ Dal punto di vista sintattico, la definizione di un *Namespace* somiglia molto a quella di una *Struttura*. e si identifica con quello di "*ambito dichiarativo con un nome*". I membri di un *namespace* sono dichiarazioni o definizioni (con eventuali inizializzazioni) di identificatori di qualunque genere (variabili, funzioni, typedef, strutture, enumeratori, altri tipi astratti qualsiasi ecc.):

<<http://www.bo.cnr.it/corsi-di-informatica/corsoCstandard/Lezioni/25Namespace.html>>

eventualmente da un altro agente Software in un circolo virtuoso di creazione e utilizzo della conoscenza.

[Meschini, 2006, p.7]

Nel contesto del Web Semantico si aprono preziose opportunità per lo sviluppo e l'utilizzo di una nuova generazione di:

- formati di metadati;
- contenuti digitali (risorse, collezioni, schemi di autorità e quelli semantici);
- tecnologie

per le Biblioteche e Archivi Digitali sfruttando le tecnologie come *Linked Data* [V.sito: *Linked Data Glossary*], che è "a recommended best practice for exposing, sharing, and connecting pieces of data, information, and knowledge on the Semantic Web using URIs and RDF" [V.sito: *Linked Data*].

La tecnologia *Linked Data* è atta a pubblicare e a connettere i dati strutturati (risorse informative e le unità che le rappresentano) sul Web. In questo modo i contenuti diventano condivisibili e ri-usabili, per cui possono essere arricchiti e validati in modo partecipato da diverse comunità di utenti" [V.sito: *Library Linked Data*].

Nell'universo bibliografico c'è un chiaro cambiamento di paradigma da formati record fissi alle ricombinabili dichiarazioni sui metadati. Le istituzioni e i sistemi bibliotecari che implementano archivi di dati bibliografici e relativi servizi dovranno necessariamente prevedere processi e strategie di fornitura dati compatibili con il framework *Linked Data* e di licenze adeguate [De Robbio, 2012, pp. 16-17].

1.2.1. Linguaggi e soluzioni tecnologiche del Web Semantico

In particolare, le tecnologie più rilevanti del Web Semantico, definiti dal *World Wide Web Consortium* (W3C) sono:

1) *XML (eXtensible Markup Language)*, una sintassi elementare per strutturare il contenuto dei documenti digitali. Attualmente, nelle tecnologie del Web semantico, XML non è una componente necessaria, considerando l'esistenza di una sintassi alternativa come *Turtle (Terse RDF Triple Language)*, uno standard *de facto* per la serializzazione dei grafi RDF.

2) *XML Schema*, un linguaggio di descrizione del contenuto XML, il cui scopo è delineare quali elementi sono permessi, quali tipi di dati sono associati ad essi e quale relazione gerarchica hanno fra loro gli elementi contenuti in un file XML.

3) *URIs (Uniform Resource Identifiers)*, indirizzi/identificatori univoci di risorse sul Web. URIs possono esprimersi tramite URN, DOI, URL accompagnati dal prefisso HTTP in modo che gli oggetti possano essere individuati da persone e da *user agent* sul Web.

4) Il già menzionato linguaggio RDF (*Resource Description Framework*), una sintassi formale di base per la codifica, lo scambio e il riutilizzo di metadati strutturati che possono descrivere la semantica di una risorsa, sia di quella reperibile sul Web (una pagina HTML, un documento XML o parte di esso), che di quella che non vi si trova direttamente (un libro, un quadro). L'utilizzo di RDF per la descrizione semantica delle risorse offre, ovviamente, numerosi benefici in molti settori del Web. Si potrebbero:

- sviluppare diverse connessioni semantiche tra le risorse digitali Web, consentendo così l'interoperabilità tra applicazioni che si scambiano informazioni sul Web, permettendo la fluente accessibilità a insiemi di dati relazionati tra loro, così da poter passare da un dato informativo all'altro senza dover preoccuparsi della provenienza diversificata degli stessi;
- lanciare motori di ricerca più efficienti in grado di basare la ricerca non solo sulle occorrenze di parole contenute nei documenti ma anche in base alla caratterizzazione semantica dei documenti stessi;

- realizzare agenti Software per il filtraggio dei contenuti di una risorsa in funzione di determinati criteri semantici impostati dall'utente, facilitando il recupero dell'informazione non solo per l'utente-uomo ma anche per l'utente – macchina..

RDF si esprime attraverso due componenti:

- 1) *RDF Model and Syntax* [V. sito] e
- 2) *RDF Schema* (RDFS o *RDF Vocabulary Description Language*) [V. sito].

Mentre la prima componente definisce il modello dei dati specificandolo attraverso la sintassi (RDF/XML, N3, Turtle, RDFa) per descrivere le risorse, la seconda permette di definire il significato e le caratteristiche delle proprietà e delle relazioni che esistono tra le risorse descritte dalla prima componente.

RDF Data Model è basato su uno *statement* costituito da una tripla del tipo:

1. SOGGETTO (una risorsa descritta mediante RDF),
2. PREDICATO (una proprietà, un attributo, una relazione definiti tra risorsa e valore tramite RDF Schema),
3. OGGETTO (un valore della proprietà)

appoggiati, nella maggior parte dei casi, dagli identificatori delle risorse URIs che puntano ad altre risorse creando un reticolo virtuoso delle informazioni correlate. Un esempio della triplicazione dei dati/contenuti informativi secondo l'articolazione (asserzione) "Soggetto-Predicato-Oggetto", ad esempio, è:

1. *la pagina* <<http://www.w3.org/DesignIssues/LinkedData.html>>
2. *è scritta dall'autore*
3. *Tim Berns Lee*.

L'approccio RDF di rappresentazione dati è strettamente correlato al nostro modo di pensare e di costruire concetti (Figura 62):



Figura 62. Esempio di componenti di uno statement di asserzioni RDF "Soggetto-Predicato-Oggetto"

L'Oggetto (Valore) (*Tim Berns Lee*) dello statement rappresentato in Figura 62, a sua volta, può essere un Soggetto (Risorsa) che sarà successivamente anche triplicata.

In questo modo si possono aprire diversi preziosi percorsi di approfondimento informativo relativi alla Risorsa di partenza, generando e arricchendo man mano il grafo RDF (Figura 63) che rappresenta nient'altro che un sapere-flusso e reticolo delle conoscenze ovvero le Ontologie con le Classi, loro proprietà e relazioni tra le Classi e proprietà per interpretare la semantica delle risorse. [V. sito: Manuale per l'interazione con gli utenti del Web culturale].

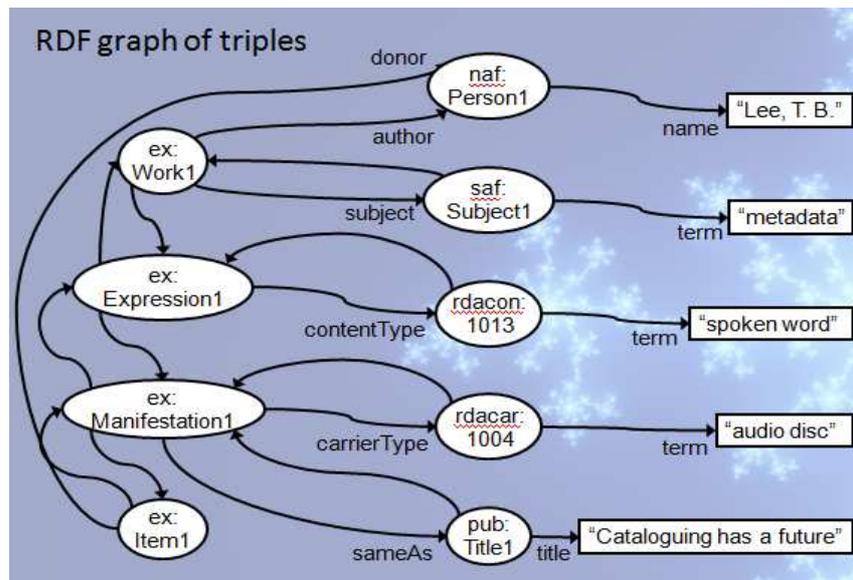


Figura 63. L'esempio dei grafi formati dalle triple *RDF* alla luce della scheda di un catalogo [Dunsire, 2012]

Nel contesto RDF, Ontologie sono i vocabolari (definizioni di Classi, istanze, proprietà) che catturano le strutture di dati di grafi RDF collegati. Nonostante l'esistenza di molti vocabolari RDF di dominio pubblico - che possono essere facilmente riutilizzati se coincidono con la semantica del dominio *target* - molti domini applicativi richiedono Ontologie specifiche che possono riflettere meglio la semantica di dati e loro relazioni.

D'altra parte, diversi gruppi di utenti lavorano in direzione di sviluppo delle specifiche RDF comuni per l'annotazione armonizzata tra diverse tipologie di risorse relazionati. Tra le recenti specifiche RDF all'avanguardia in questa direzione sono, ad esempio:

- *ANNOTATION ONTOLOGY* e
- *OPEN ANNOTATION MODEL* conforme all'architettura del *World Wide Web* e all'iniziativa *Linked Data* [V. sito: *Open Annotation Community Group* (OACG)].

Open Annotation Model può essere usato anche per annotare le risorse informative scientifiche tramite le procedure di *Open Peer Review*, la creazione di Nano-pubblicazioni [V. APPENDICE B] e altro ancora.

Sopra i costrutti dei:

1) *RDF Model and Syntax* [V. sito] - che consente unicamente indicare predicati binari (valore-attributo) e

2) *RDF Schema* RDF - che consente di stabilire gerarchie di Classi e Proprietà e di imporre vincoli per dominio e codominio - il W3C ha definito *Web Ontology Language* (OWL) [V. sito], il linguaggio OWL permette di *rappresentare* esplicitamente il *significato* (semantica) delle strutture informative che popolano il Web mediante vocabolari (Ontologie) web formali.

In particolare, l'OWL esprime non solo Proprietà (*properties*), Classi (*classes*), Istanze delle classi (*individuals*) e Relazioni (di disgiunzione, di cardinalità, di uguaglianza) tra le classi, ma anche esporta tale conoscenza a diverse applicazioni, contribuendo maggiormente alla loro interoperabilità. L'espressività di OWL è costituita da tre sottolinguaggi:

- *OWL Lite*,
- *OWL DL* (*Description Logic*),
- *OWL Full*,

ciascuno dei quali riproduce un'estensione del suo modello precursore più semplice (non viceversa). *OWL Lite* rappresenta un percorso di migrazione più rapida per Thesauri e altre tassonomie gerarchiche e sistemi inferenziali, supportando, però, solo i valori di cardinalità a

zero o uno, ed escludendo alcuni costrutti che esprimono relazioni tra le Classi (ad esempio disgiunzione, unione, istanziazione).

OWL DL è abbastanza espressivo per trattare tutti i costrutti OWL, tuttavia assegna solo alcuni vincoli sul loro uso (ad esempio una Classe può essere Sottoclasse di un'altra classe, ma non un'istanza di una classe; non si possono esprimere restrizioni di cardinalità per le proprietà transitive).

OWL Full rende la massima espressività e la libertà sintattica dei costrutti RDFS, ma è difficilmente implementabile.

6) La tecnologia *Simple Knowledge Organization System (SKOS)* [V. sito], basata su RDF e RDFS. SKOS appoggia la strutturazione e la pubblicazione dei sistemi per l'organizzazione della conoscenza (*Knowledge Organisation Systems, KOS*) quali Thesauri, Schemi di classificazione, Soggettari, Schemi di intestazione per soggetto, e altri vocabolari controllati sul Web.

Ad esempio, il NUOVO SOGGETTARIO della Biblioteca Nazionale di Firenze [Bergamin, Lucarelli, 2013] è stato pubblicato attraverso i formalismi SKOS.

Altri KOS resi disponibili tramite i formalismi SKOS si possono consultare nel Rapporto *Library Linked Data Incubator Group: Datasets, Value Vocabularies* [V. sito].

7) *SPARQL (Simple Protocol and RDF Query Language)* [V. sito], un linguaggio di interrogazione dei grafi e sottografi RDF, il cui *output* sarà l'estrazione delle informazioni dalle basi di conoscenza distribuite sul Web.

Le tecnologie appena descritte rappresentano strati sovrapposti di linguaggi, ognuno dei quali usa o estende gli strati precedenti. Il collocamento di queste tecnologie può essere osservato nella Figura 64 [V. sito: *Semantic Web Architecture*].

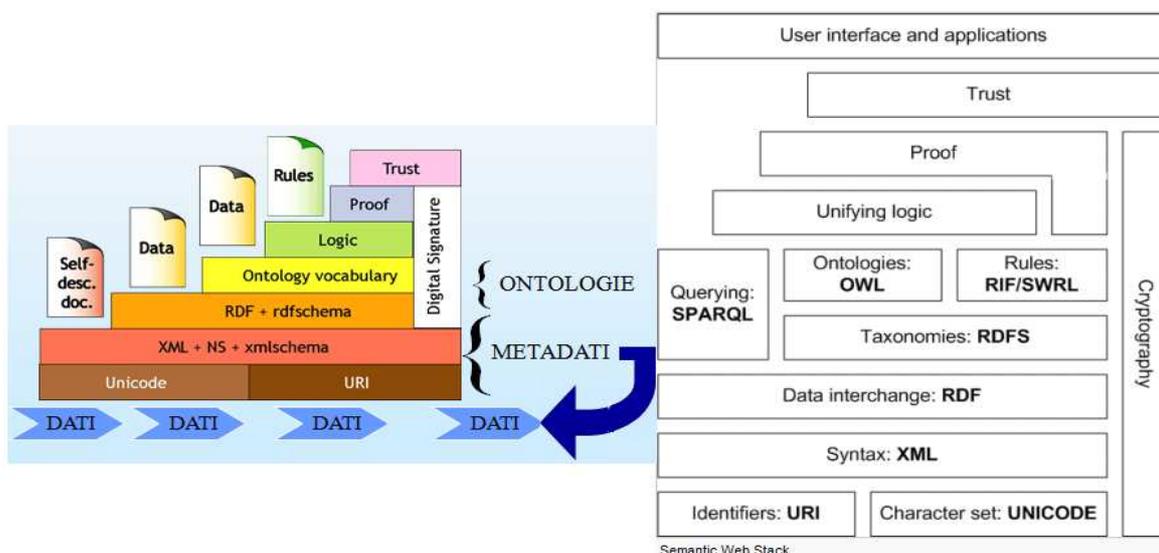


Figura 64. Le tecnologie del Web Semantico

La rappresentazione nella Figura 64 può essere interpretata come segue:

1) fino al livello delle Ontologie (metadati strutturati secondo un ordine logico, relazionati attraverso i concetti di un dominio di conoscenza) non vi è inferenza, ma solo rappresentazione della conoscenza.

2) Sopra questo livello si colloca il livello Logico che, tramite certi linguaggi (come SWRL, RuleML) e ragionamenti inferenziali automatici, estrae dalle asserzioni (*statements*) Web una nuova conoscenza (*machine understandable*), che a sua volta sarà automaticamente integrata e riutilizzata dalle applicazioni.

3) Per raggiungere la piena potenzialità, le informazioni del Web Semantico devono essere approvate (*Proof*) e, di conseguenza, affidabili (*Trusted*).

Uno dei problemi fondamentali del Web Semantico è quello di rendere disponibili, integrabili e interoperabili le varie tipologie di dati sul Web. Tecnicamente, questo obiettivo può essere raggiunto attraverso:

- opportune tecnologie capaci di operare la conversione da diversi formati di dati e database in RDF [V. sito: *W3C RDB2RDF Incubator Group*]
Tra i maggiori servizi che si prestano alla trasformazione dei dati strutturati in formati RDF/OWL possono essere elencati: *Web services links & resources*, *SemWev*, *Beckett*, *SIMILE (RDF crosswalks)*, *Semantic Bank*, *D2R Server* (per l'esposizione di database relazionali in RDF);
- *markup* RDF automatizzato dei siti (ad esempio attraverso le tecnologie POWDER, RDFa GRDDL, R2RML, RIF, Drupal) [De Robbio, 2011].

Programmare l'interoperabilità dei sistemi e applicazioni del Web Semantico è possibile anche tramite i servizi di *tagging* semantico *API (Application Programming Interface)*.

API è "un modo per i programmi di comunicare tra loro. [Lo] Si può comprendere nei termini di come un programmatore invia istruzioni tra programmi diversi" [V. sito: *Open Data Handbook*]. Sono esempi di API: OpenCalais e Zamanta. Questi sono utiliper trasformare blocchi di testi non strutturati in entità principali, *topics*, relazioni, e per eseguire l'estrazione di *keywords*, *tagging* automatico, e la disambiguazione di entità e concetti, che possono fungere da output in RDF.

1.2.1.1. *Linked (Open) Data (Web 3.0): il valore delle relazioni per la posizione partecipativa dei contenuti sul web*

Il termine *Linked Data (LD)* è stato coniato da Tim Berners-Lee (il creatore del *World Wide Web*) nel 2006, in connessione all'ambito del Web Semantico. Nel suo famoso scritto introduttivo sui *Linked Data* Tim Berners-Lee [2006] ha ribadito che il Web Semantico non significa soltanto pubblicare i dati sul Web. Il Web Semantico si riferisce alla produzione dei link, delle connessioni, così che utenti, sistemi e applicazioni Web possano esplorare:

- non solo i documenti ipertestuali del Web tradizionale descritti in HTML ma anche
- il *Web di dati* ovvero di entità informative o *dataset* (pacchetti di contenuti) digitali di diverso tipo e formato descritti in RDF, partendo dalla scoperta dei quali è possibile trovare le altre entità connesse.

Per appoggiare la crescita del Web, la tecnologia *Linked Data (LD)*:

- usa gli identificatori URIs per identificare le entità informative;
- si riferisce agli URIs basati sul protocollo del Web HTTP, affinché gli agenti Web possano ricercare e interpretare i medesimi identificatori sul Web;
- impiega la tecnologia RDF per esprimere, strutturare e pubblicare i dati sul Web;
- si serve di altri URIs_HTTP connettendoli agli URIs di partenza, per favorire la scoperta di ulteriori entità informative sul Web.

Con l'approccio appena esposto, la tecnologia LD ha come l'obiettivo principale di estendere il Web attuale ad uno spazio globale di dati (o *Web of data*, *Web 3.0*) alla portata di tutti [Heath, Christian, 2011]. *Web of data* può essere concepito come un *database* globale contenente i dati strutturati e linkabili, come se fossero pagine Web:

si costruisce così un reticolo di dati collegati (*Linked Data*, appunto) appartenenti a un dominio (che costituisce il contesto di partenza), collegato a sua volta ad altri set di dati esterni, ovvero fuori dal dominio, in un contesto di relazioni sempre più estese [Guerrini, Possemato, 2012, .p.7].

I *dataset* del *Web of data* possono provenire dal dominio di dati:

1) Privato (Dati Chiusi):

While, to date, it is the case that *Linked Data* has been demonstrated using public Web data and many desire to expose more through the *Open Data* movement, there is nothing preventing *private*, *proprietary* or *subscription data* from being *Linked Data* [V. sito: *What is Linked Data?*];

2) Pubblico - che contiene Dati Aperti attraverso le Licenze Aperte [De Robbio, 2012, p.7; V. sito: *Open definition*].

Collegando i Dati Aperti tramite la tecnologia LD, si dà il luogo alla creazione di *Linked Open Data* (LOD) [*Linked Open Data. What is it?*], che permettono di avere una maggiore utilità di contenuti (accesso, consultazione, arricchimento, ri-uso di dati).

I dati rappresentati sul Web tramite il RDF (che descrive i concetti e le relazioni su di essi attraverso l'introduzione di triple Soggetto-Predicato-Oggetto) vengono interrogati tramite il linguaggio *SPARQL* (*Protocol and RDF Query Language*) [V. sito] (Figura 65).

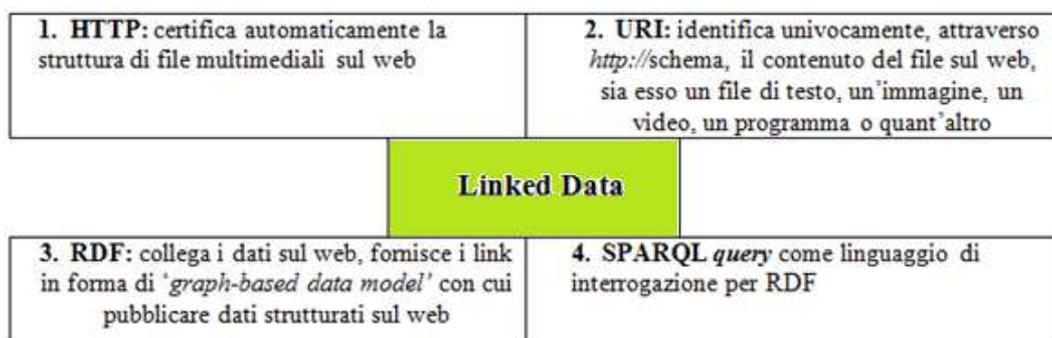


Figura 65. Componenti della tecnologia *Linked Data*

SPARQL è uno degli elementi chiave delle tecnologie legate al paradigma del Web Semantico, e consente di estrarre informazioni (tramite la costruzione di *Query* basate su triple patterns, congiunzioni logiche, disgiunzioni logiche, e pattern opzionali) dalle basi di conoscenza distribuite sul Web, e quindi anche informazioni pubblicate con la tecnologia *Linked Data*.

Riguardo al rapporto tra *Linked Data*, *Open Data* e *Web Semantico* è interessante la definizione data da Paul Walk [V. sito: *Linked, open, semantic?*]:

- 1) i dati possono essere aperti, ma non linkati;
- 2) i dati possono essere linkati, ma non essere aperti;
- 3) i dati che *siano sia aperti sia linkati* sono assolutamente realizzabili e riutilizzabili;
- 4) il Web Semantico può funzionare solo con dati che siano sia aperti sia linkati (*più dati aperti sono linkati sul Web con un significato ben definito, più le applicazioni Web saranno performanti*).

Nel Web di Dati – rappresentato graficamente da *Linking Open Data Cloud Diagram* [V. sito] (Figura 66) - diversi *dataset* forniti da una varietà di comunità di utenti sono resi integrabili, condivisibili e ri-usabili, offrendo un accesso flessibile [Cyganiak, et al., 2008] e multifocale ai contenuti correlati da parte di diverse applicazioni Web che, a loro volta, acquistano un nuovo livello di integrazione e di interoperabilità, arricchendo ulteriormente la rete di collegamenti sul Web [Cyganiak et al., 2008; V. sito: *Dublin Core. Building blocks for interoperability*].

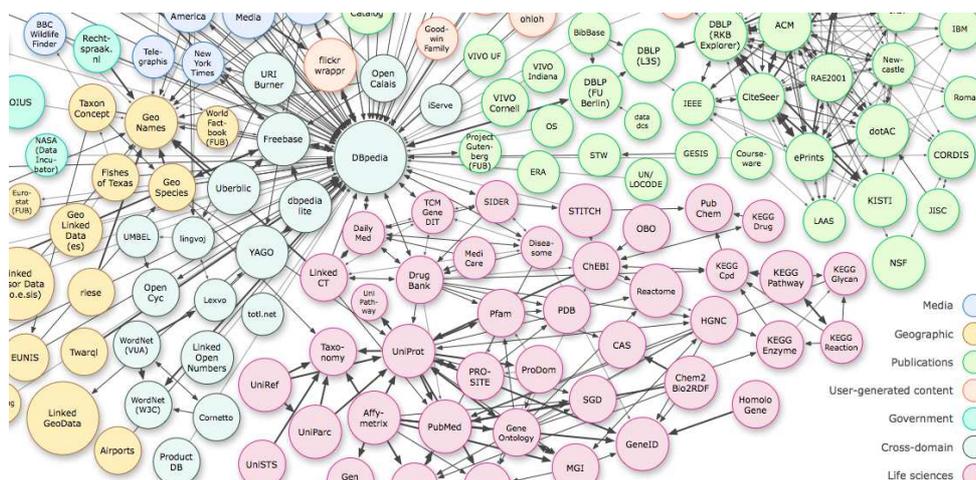


Figura 66. Rappresentazione delle connessioni tra diversi dataset nel *Linked Open Data (LOD) Cloud Diagram*

Il diagramma *LOD Cloud* è stato sviluppato nel 2007 entro il progetto *Linking Open Data* della *SWEO community* [V. sito] per visualizzare graficamente la pubblicazione e i collegamenti dei dati dai domini privati e quelli pubblici (dati pubblicati sotto le Licenze aperte come, ad esempio, la CC0, PDDL, ODC-BY e non condizionati da licenze commerciali o restrizioni d'uso) sul Web tramite *Linked Data*.

I dati facenti parte della *LOD Cloud* (in continua crescita) rappresentano un immenso database di entità relazionati che, in caso di dati aperti, possono essere liberamente incrociati da terze parti, con la possibilità di generare valore aggiunto per i servizi personalizzati, magari in base alle esigenze di una tipologia specifica di utenti.

Mentre nel 2007 il diagramma LOD contava più di 2 miliardi di terne RDF provenienti da diversi database Web (le prime e più aggiornate sono DBpedia e Freebase) [Sandhaus, Rob, 2009] nel 2013 il *LOD Cloud* ha raggiunto oltre di 55 miliardi di terne RDF.

La crescente richiesta da parte degli utenti di servizi che consentono un elevato grado di interconnessione tra i dati provenienti da diverse fonti ha portato allo sviluppo collaborativo e alla promozione di soluzioni pratiche per LOD in diversi settori, puntando ai fattori di maggiore interesse come:

- l'interoperabilità semantica,
- la migliore fruibilità,
- l'ampio riutilizzo di entità informative e
- la loro libera, aperta disponibilità per l'utente.

L'importanza di LOD viene ampiamente riconosciuta anche per:

- l'annotazione semantica [Rusu, et al, 2011],
- l'arricchimento dei dataset web attraverso modelli concettuali e *authority data* resi disponibili attraverso RDF [Gordon, 2010] e *Linked Open Data* [V. siti: *Library Linked Data Incubator Group: Datasets, Value Vocabularies; Use Case Authority Data Enrichment*]

per l'implementazione di *high-level web-scale services* per la ricerca.

In prospettiva dello sviluppo delle strategie di pubblicazione delle entità informative come *Linked Data* e con il taglio tematico per i dati bibliografici, è interessante prendere visione delle seguenti iniziative:

- *Archives & Museums (LODLAM) Group*,
- *IFLA'S Semantic Web Special Interest Group*,
- *Library Linked Data Group*,
- *Learning Linked Data*,

- *Linked Open Vocabularies*,
- *LOD2Project*,
- *Linked Open Data-enabled Bibliographical Data (LODE-BD)*

nonché delle relazioni di alcuni recenti eventi come:

- *Global Interoperability and Linked Data in Libraries*,
- *2nd Linked Open Data Conference* [V. siti], e altro ancora.

Ormai diverse istituzioni hanno pubblicato le loro entità (risorse, collezioni, vocabolari, metadati) in modalità *Linked Open Data*, per favorirne la consultazione, lo scambio, il riuso e l'arricchimento sul Web. Tra tali istituzioni ci sono:

- 1) l'Università di Münster che coordina il progetto *The Linked Open Data University of Münster (LODUM)* entro cui è stata sviluppata un'infrastruttura per la pubblicazione di metadati descrittivi i lavori scientifici e quelli didattici prodotti entro la medesima Università in modalità LOD;
- 2) la *Library of Congress*, che ha reso il proprio servizio *Authorities and Vocabularies* [V. sito: *Library of Congress Authorities*] in modalità LOD. Tale servizio dispone di un sistema terminologico *Semantic Web Tools for Libraries*. L'attenzione verso LOD è ben espressa nell'iniziativa della Biblioteca *A Bibliographic Framework for the Digital Age* [V. sito];
- 3) la *German National Library (DNB)* che, in cooperazione con la *German Serials Database (Zeitschriftendatenbank – ZDB)*, ha creato *Linked Data Service* [V. sito: *German National Library goes LOD & publishes National Bibliography*].

Questo servizio esegue l'*Authority Data Linking* dei record bibliografici (schede di metadati) disponibili (inclusi *Authority Data* proposti dalle grandi agenzie come la *Library of Congress* e la *British Library*) con le entità fornite dai dataset *Linked Open Data - DBpedia* e *VIAF (Virtual International Authority File)* [Keßler, 2010].

Attualmente ci sono più di 1.6 milioni di record bibliografici con l'identificatore per titoli seriali ZDB-ID che, a differenza di ISSN, distingue le versioni a stampa da quelle elettroniche [V. sito: *Zeitschriftendatenbank*];

- 4) la *Hungarian National Library*, che pubblica i propri *dati bibliografici* e quelli di *autorità* come *Linked Data*, arricchendoli ulteriormente attraverso schemi semantici come RDFDC, FOAF e SKOS [V. sito: *Hungarian National Library OPAC*];
- 5) la *British National Library (BNB)* che, oltre a fornire i propri dati con RDF download, li pubblica come LD collegandoli ai dataset *Linked Open Data* come VIAF, LCSH (*Library of Congress Subject Headings*), Lexvo, GeoNames, MARC country, Dewey.info, RDF Book Mashup [Wilson, 2010];
- 6) la *Bibliothèque Nationale de France* che, tramite il Progetto *Data.bnf.fr* ha reso i dati bibliografici relativi a opere e autori maggiormente integrati e fruibili sul Web attraverso le terne RDF, seguendo i principi di LOD [V. sito: *Data.bnf.fr: an overall presentation*];
- 7) la *Biblioteca Nacional de España (BNE)* la quale, in collaborazione con l'*Ontology Engineering Group (OEG)* dell'Università Politecnica di Madrid (UPM), ha lanciato il Progetto aperto *Linked Data at the BNE* [V. sito: *Datos enlazados en la BNE*]. Attraverso il tool *MARiMba* [V. sito] circa 2,5 milioni di record bibliografici (metadati su monografie antiche e moderne, registrazioni sonore) e 4 milioni di *authority record* della BNE sono stati trasformati in RDF. Il processo di trasformazione ha generato circa 58 milioni di terne RDF e 600 collegamenti (owl:sameAs) ad altri set di dati come DBpedia o VIAF [V. sito: *Datos.Bne.Es*], abilitando così i dataset della Biblioteca come *Linked Data* sotto la Licenze Aperte CC0;
- 8) la Biblioteca di *North Rhine-Westphalia*, che ha lanciato il servizio LOD - LOBID.ORG [V. sito: *Data*], dove LOBID sta per *Linking Open Bibliographic Data* ovvero dati bibliografici linkati entro e all'esterno della Biblioteca con una varietà di dati associati;

9) l'*Online Computer Library Center (OCLC)*, che ha reso disponibile oltre 1 milione di file delle più importanti opere bibliografiche, contenute nel catalogo digitale globale *WorldCat*, in modalità di download *Linked Data*, sotto la licenza *ODC-BY*. La serializzazione *RDF* è basata sull'Ontologia web *SCHEMA.ORG* e sulle estensioni a tale ontologia dei dati bibliografici:

this is an important step for Libraries and Linked Data [...] Organizations wishing to develop LINKED DATA SERVICES can experiment with this data set before going into full development. They'll also be able to stress-test new services using a very large and important set of up-to-date, LINKED LIBRARY DATA. We are really interested to see what people will do with this data [...] OCLC expects that the file will be useful as a source of raw data. Information about works, authors and publishers can be dissected and recombined in this format much more easily
[V. sito: *OCLC provides downloadable Linked Data file*];

10) La *Bayerische Staatsbibliothek* che, insieme ad altre Biblioteche [V. sito: *Europeana Libraries*], sta provvedendo all'allineamento dei propri metadati bibliografici con l'*Europeana Data Model (EDM)* [Doerr et al., 2010] e alla loro pubblicazione come *LOD* per rendere i medesimi dati apertamente consultabili, ri-utilizzabili e maggiormente interoperabili sul Web.

Seguendo gli sviluppi del progetto *Linked Data*, nel 2011 il gruppo di lavoro *W3C Library Linked Data (LLDXG)* [V. sito: *Library science talk*] – con lo scopo di sviluppare una standardizzazione condivisa del Web Semantico all'interno della comunità Bibliotecaria – ha pubblicato il rapporto "*Library Linked Data*" (LLD), in cui vengono spiegate modalità come possono essere pubblicati in modalità *Linked (Open) Data*:

- dati bibliografici (metadati);
 - dati di autorità (*Authority data*; ad es. dati di vocabolari controllati come Thesauri, Classificazioni);
 - risorse bibliografiche e collezioni di Archivi;
 - citazioni e altre entità facenti parte di Biblioteche e Archivi digitali,
- per permettere il loro ri-uso (inclusa la creazione di un'ulteriore varietà di relazioni) sul Web da parte di diversi agenti. La versione grafica dei pacchetti di *Library Linked Data (LLD)* - pubblicati sul Web Semantico e descritti nel Rapporto *Library Linked Data* – vengono presentati in Figura 67.

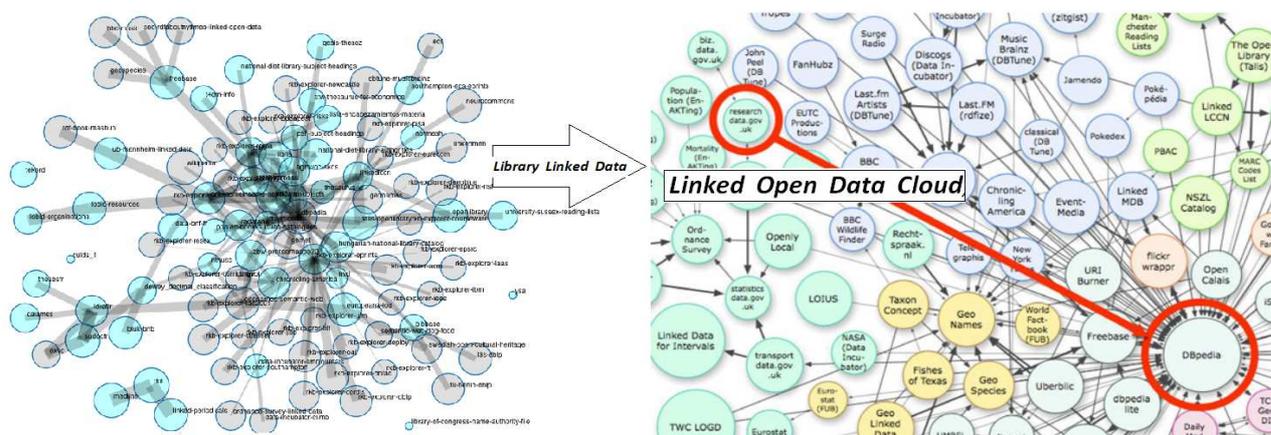


Figura 67. *Library Linked Data (LLD)*: verso l'integrazione nello spazio globale dei dati interconnessi *Linked Open Data Cloud*

I *dataset* come Vocabolari Controllati e Schemi di Classificazione presentati nel Rapporto *Library Linked Data* possono diventare maggiori fornitori di *authoritative datasets* nello spazio del Web Semantico. Particolarmente, gli *authoritative datasets* comprendono le entità in forma normalizzata come:

- nomi di agenti (persone, organizzazioni);
- valori che identificano classi, soggetti/temi, luoghi geografici.

Con riferimento all'ambiente digitale, è evidente l'importanza di qualificare la descrizione delle risorse Web con gli strumenti di *Authority Control* in particolare, per quanto riguarda le descrizioni di Persone fisiche ed Enti collettivi, e l'individuazione del ruolo che questi soggetti svolgono nella creazione, produzione, diffusione e conservazione dei singoli documenti.

Inoltre, integrandosi nello spazio di una diversità di pacchetti di dati collegati della *LOD Cloud*, diversi *dataset* bibliografici, a seconda delle relazioni sviluppate, possono arricchire i propri profili semantici. Generando così ricchi *cluster* (gruppi) di entità, di proprietà descrittive e normalizzatrici, sarà anche possibile elaborare varie *Query sofisticate* per interrogare le fonti bibliografiche e altre risorse sul Web. In merito ai ricchi *cluster* di entità *Linked (Open) Data* si vedono, oltre agli esempi di esperienze già forniti in precedenza, anche:

1) il Portale *WordCat* [V. sito] - il più grande Catalogo online di risorse bibliografiche provenienti da diverse Biblioteche e Archivi a livello globale e collegate tra loro tramite le tecniche di *Linked Data*;

2) *OPEN LIBRARY* [V. sito] - una Biblioteca online universale e aperta che intende costruire un catalogo digitale che annoveri ogni libro e ogni pubblicazione esistente al mondo, attorno ai quali possano gravitare una costellazione di dati prodotti dagli utenti, tramite la tecnologia *Linked Data*;

3) *EUROPEANA* [V. sito] - una Biblioteca digitale europea che riunisce contributi già digitalizzati da diverse istituzioni dei 27 paesi membri dell'Unione Europea e li presenta in modalità *Linked Open Data* (tramite metadati aperti e collegati);

4) *TEL: THE EUROPEAN LIBRARY* - il sistema che permette di consultare i Cataloghi online di 48 Biblioteche nazionali incluso *European Research Libraries*. Il Portale TEL offre agli utenti registrati la possibilità di ricerca tramite i Vocabolari di autorità e Schemi semantici disponibili in modalità *Linked Open Data (LOD)* connettendo tramite essi la conoscenza di diverse collezioni e risorse, e contribuendo quindi a *LOD Content Negotiation* tra esse [V. sito]. La descrizione di *dataset* (risorse, collezioni) della TEL seguono il modello *Europeana Data Model/EDM* [V. sito].

5) le infrastrutture digitali "Burckhardtsource.org" (una biblioteca digitale creata in seno al progetto "*The European correspondence to Jacob Burckhardt*" [Di Donato, Müller, 2013] e *VOA3R Open Access Repository*) [V. sito; Subirats, Zeng, Keizer, 2011] che permettono l'arricchimento, l'interconnessione e la ricerca di contenuti di valore culturale e scientifico con il supporto delle strategie *Linked Data*.

6) *Public Library of Veroia in Web 3.0* [V. sito].

Nonostante gli indiscutibili vantaggi di LLD nel mondo del Web Semantico e per i numerosi progetti che promuovono i metodi innovativi di pubblicazione di dati (contenuti digitali), sarà necessario un *coordinamento specifico, accurato e deciso*, al fine di contrastare il fenomeno dell'implementazione irregolare di tecnologie e principi LD da parte di diverse comunità di utenti Web.

Tra i principali *ostacoli* alla *pubblicazione dei dati con i meccanismi Linked Data* vi è ancora la riluttanza, da parte di diverse comunità, nell'adottare nuove piattaforme Software, così come la mancanza del personale con adeguate conoscenze e competenze per utilizzare queste piattaforme, prima e dopo la loro implementazione [Jackson, 2009].

Se i bibliotecari, esperti in catalogazione, elaborazione e gestione di metadati si sentissero sopraffatti dalla prospettiva di produzione di metadati semantici e di ontologie, allora lo

sarebbero ancora di più altre comunità operanti sul Web. Inoltre, esistono tensioni tra *Open Standards*, e *Open Publishing* richiesti per il Web Semantico e i metodi esistenti per il supporto e il mantenimento delle infrastrutture di gestione dati.

Da un punto di vista positivo, si potrebbe imitare le esperienze di alcune istituzioni (delle quali abbiamo dato un accenno in precedenza) che pubblicano i loro dati bibliografici con la tecnologia LD.

Per contrastare le difficoltà di integrazione di dati bibliografici nel Web Semantico, il Rapporto *Library Linked Data* ha fornito quattro principali raccomandazioni, con cui:

1) invita le comunità bibliotecarie a individuare pacchetti di dati (inclusi Thesauri e altri vocabolari controllati) da pubblicare in modalità LOD e in linguaggi standard del Web semantico (ad esempio SKOS), così da aprire le porte alla creazione di nuovi e potenzialmente illimitati servizi; invita, dunque, a diffondere quanto più possibile le proprie esperienze riguardo al *Free, Open Data* [V. sito: *Linked Open Data Italia*], attraverso Licenze Aperte e formati non proprietari. Nella relazione vengono individuate tre potenziali categorie di dati bibliografici da pubblicare in LD e gestire con gli appositi *namespace*⁴:

a) *dataset* formanti le terne su risorse bibliografiche, persone, organizzazioni e soggetti connessi;

b) vocabolari controllati da cui possono derivare le terne formate da valori, proprietà e attributi per *Authority File*;

c) formati e vocabolari (schemi) di metadati che generano le terne descrittivi attributi e relazioni usati entro *dataset* bibliografici e vocabolari controllati;

2) si appella ai gruppi di lavoro degli enti che partecipano nella produzione degli standard (es. *IFLA: Namespace Technical Group reposting to Committe on Standards*; *JSC for Development of RDA*; *DCMI/RDA Task Group: Dublin Core Metadata Initiative*; *DCMI Bibliographic Metadata Task Group: DCMI Vocabulary Management Community*) per supportare la partecipazione delle Biblioteche nella standardizzazione del Web Semantico e nello sviluppo di Standard per i dati bibliografici compatibili con *Linked Data*, e per disseminare *best practices* sulla modellazione di *Library Lined Data*;

3) si appella, poi, per influenzare i *vendor* (le aziende commerciali fornitrici di risorse elettroniche e database online) affinché vengano considerati e adottati questi Standard;

4) si rivolge agli sviluppatori dei sistemi informativi digitali affinché essi:

a) progettino *high-level services* di gestione dei dati basati sulle capacità del Web semantico e di *Linked Data*,

b) creino URIs (anche in modo sperimentale) per poter puntare agli elementi di *dataset* bibliografici e a dati prodotti da altre comunità in modo da definire un contesto,

c) sviluppino politiche per la gestione dei vocabolari RDF e di relativi URI;

d) esprimano i *dataset* bibliografici in termini di loro riuso o mappatura in collegamento a esistenti vocabolari pubblicati in *Linked Data* e adeguati al contesto di riferimento;

4) invita le comunità di Bibliotecari e archivisti a preservare i loro *dataset* e Vocabolari Controllati in set di elementi *Linked Data* e ad applicare i loro *know-how*, riferiti alla gestione dati e alla conservazione a lungo termine, ai *dataset* pubblicati in *Linked Data* così che le terne di *Linked Data* diventino immutabili, gli identificatori univoci URIs rimangano permanenti e nomi, titoli, intestazioni per soggetto permangano stabili nel tempo. Con il progredire delle tecnologie del Web semantico, le biblioteche e le istituzioni archivistiche avranno una buona opportunità per arricchire i loro attuali ruoli di gestori dell'universo bibliografico con la qualifica di autorità garanti della conservazione a lungo termine dei *dataset* in *Linked Data* relativi al patrimonio scientifico e storico- culturale: "*Libraries and librarians have a very high "trust"*

⁴ Ad esempio, entro la *British National Bibliography* (BNB) pubblicata come *Linked Data* dalla British Library vengono gestiti diversi *namespace* (MARC, VIAF, ISBD, DC, LCSH, Lexvo, GeoNames): <<http://thedatahub.org/it/dataset/bluk-bnb>>.

factor with the general public, and we are in a position to raise the general quality of the Semantic Web by ensuring Library data competing with data from the widest range of sources” [Dunsire, online, non impaginato].

1.2.1.2. Topic Maps: una tecnologia per organizzare e rappresentare la conoscenza

Le *Topic Maps (TM)* sono una tecnologia del Web Semantico che rappresenta un modello concettuale per l'organizzazione e la rappresentazione della conoscenza sul Web.

Le TM sono state approvate come standard ISO IEC 13250 atto alla rappresentazione di strutture associative attraverso lo schema di codifica delle *Topic Maps XTM* e *Linear Topic Map Notation (LTM)*, specificati nello stesso standard. XTM è una specializzazione del linguaggio XML adatto per la *formattazione* di informazioni a struttura associativa quali mappe concettuali, reti semantiche, basi di conoscenza e più in generale strutture logiche e cognitive. Intanto,

le *Topic Maps* non sono né mappe di navigazione, né mappe grafiche, né mappe mentali o concettuali [...] non devono essere identificate necessariamente con *eXtensible Markup Language (XML): XML Topic Maps (XTM)* è solo uno dei possibili formati di serializzazione, ma una *Topic Map* può essere immagazzinata anche all'interno di un Database Relazionale o di altri linguaggi di serializzazione [Vassallo, 2010, p.173].

Come i “*concetti*” sono gli elementi primitivi di ogni *mappa concettuale* così i TOPIC lo sono di ogni *Topic Map*. Un *Topic* può essere:

- un concetto astratto,
- un oggetto reale,
- un evento,
- un nome,

in altre parole, qualsiasi cosa che possa rappresentare oggetto di un discorso o di una dichiarazione (ad es. sono TOPIC “*Roma*”, “*Aristotele*”, “*Linguistica*”, “*Opera*”). Quindi,

Il TOPIC è la rappresentazione univoca di un qualunque “SOGGETTO”, intendendo per soggetto qualunque elemento del discorso su cui l'autore della *Topic Map* intenda parlare. TOPICS e SUBJECTS sono dunque gli ideali componenti di un TRIANGOLO SEMIOTICO: rappresentano rispettivamente il SIMBOLO e il REFERENTE. Tra il *Topic* e il *Subject* esiste una relazione uno a uno, in cui ogni Soggetto viene rappresentato da un solo *Topic* e ogni *Topic* può rappresentare un unico Soggetto. Questo principio (detto *collocation objective*, “collocazione oggettiva”) assicura che tutto ciò che si conosce all'interno di un dato Sistema) su un *determinato Soggetto* sia accessibile univocamente da un *punto d'accesso*, dato che quel Soggetto è rappresentato da solo un *Topic* [ivi, p.173].

La flessibilità della tecnologia *TM*, evolutasi dal contesto di creazione e fusione di TASSONOMIE e

di INDICI ANALITICI di manuali tecnici entro cui era nata con il nome di *Navigation Maps*, offre alcuni strumenti che possono renderla appetibile anche per la GESTIONE e FUSIONE di METADATI descrittivi differenti. Il modello di dati delle TOPIC MAPS definisce infatti alcuni elementi (costrutti nel lessico delle *Topic Maps*) che combinati fra di loro permettono di ESPRIMERE e CODIFICARE una GAMMA illimitata di DESCRIZIONI e RELAZIONI e che, per tornare allo scopo originario della tecnologia, permettono di gestire agevolmente INDICI MULTIPLI fondendoli e disambiguando i differenti termini [ivi, p.174].

TASSONOMIA e la CLASSIFICAZIONE GERARCHICA di concetti; principio stesso della classificazione. Praticamente tutti i CONCETTI, gli OGGETTI animati e non, i LUOGHI e gli EVENTI possono essere classificati seguendo uno schema tassonomico. Per CLASSIFICAZIONE si intende la DESCRIZIONE e la COLLOCAZIONE in un Sistema Tassonomico di una ENTITÀ; per determinazione si intende il riconoscimento o l'identificazione di un soggetto.

Le TM - essendo un'Ontologia che opera sui contenuti Web - consentono più ricca integrazione e l'interoperabilità dei dati tra diverse comunità di informazione in rete nel Web Semantico (Figura 68).

Le TM propongono una architettura alternativa ai linguaggi ontologici RDF/S-OWL prefigurando scenari di utilizzo differenti.

A differenza di RDF - che consente solo relazioni binarie - le associazioni (*relazioni*) predefinite di TM possono essere *n-arie*, cui costrutti permettono di esprimere:

- rapporti dipendenze "Classe/sotto-classe";
- rapporti "Tipologia/Istanza";
- direttamente la *tipologia* di un *Topic* e
- il *Ruolo* che esso svolge in una *Associazione*.

Tramite le relazioni *n-arie*, una *Topic Map* può formalizzare una *rete concettuale* di *Topic*, *Ruoli*, *Associazioni*, collegata ad un insieme delle risorse intrinseche alle medesime relazioni.

Una delle particolarità delle *TM* è di avvicinarsi al modo di pensiero associativo umano. Questo differenzia le *TM* da altri strumenti relativi al Web Semantico (ad esempio, RDF) che non risultano *human readable* [Pepper, 2009].

A titolo esemplificativo, possiamo concettualizzare attraverso *TM* un semplice enunciato:

Aristotele - è l'Autore di opere - "Poetica" e "Retorica".

Si avranno quindi i *Topic*: "Aristotele" - "Autore" - "Opera" - "Poetica" - "Retorica".

È chiaro che tra "Autore" e "Opera" c'è un'evidente Relazione bidirezionale: "un Autore scrive Opere" vs. "le Opere sono scritte da un Autore". Questa relazione è anche un *Topic* (Figura 69).

A livello più materiale, i *Topic* si manifestano nelle *Occorrenze* chiamate delle risorse "tangibili".

Per il *Topic* "Retorica" un'Occorrenza può essere:

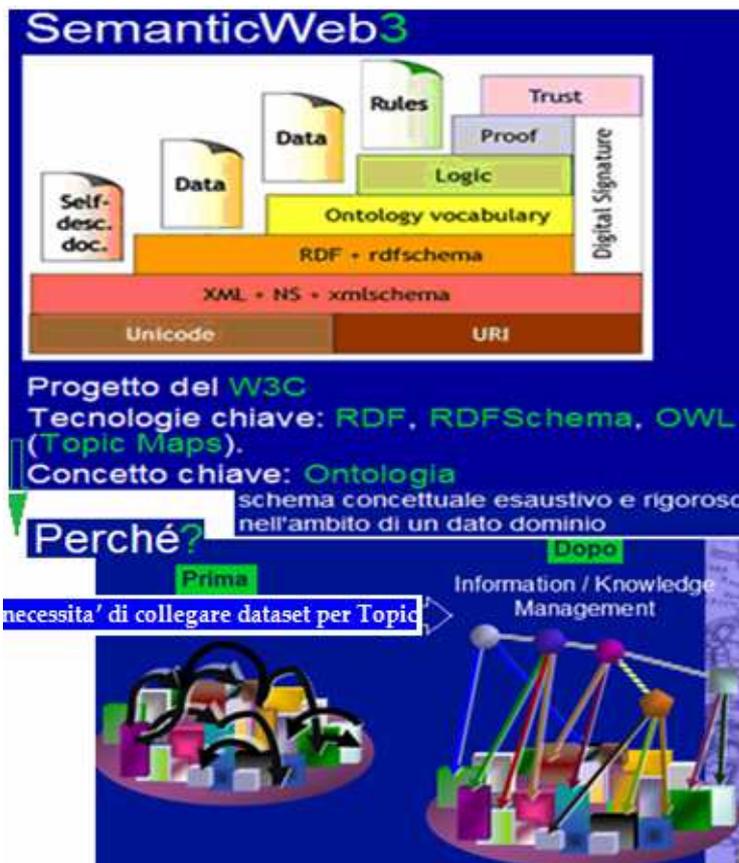
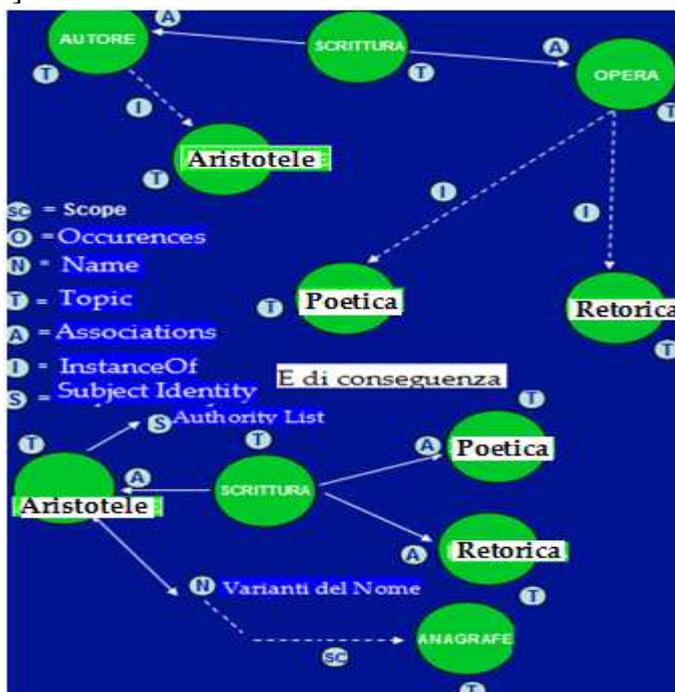


Figura 68. *Topic Maps* nei costrutti del Web Semantico



- l'edizione elettronica su *LiberLiber* o quella su Biblioteca Italiana,

- l'edizione multimediale su CD-Kom de L'Espresso,

- la versione a stampa a cura di Armando Plebe e edita da Mondadori.

Ogni Occorrenza:

- sarà identificabile tramite un identificatore *Uniform Resource Identifier* basato sul protocollo Web HTTP;

- avrà una sua tipologia (file HTML, XML, Cd-Rom, libro a stampa), a sua volta rappresentata da un *Topic* (Figura 69).

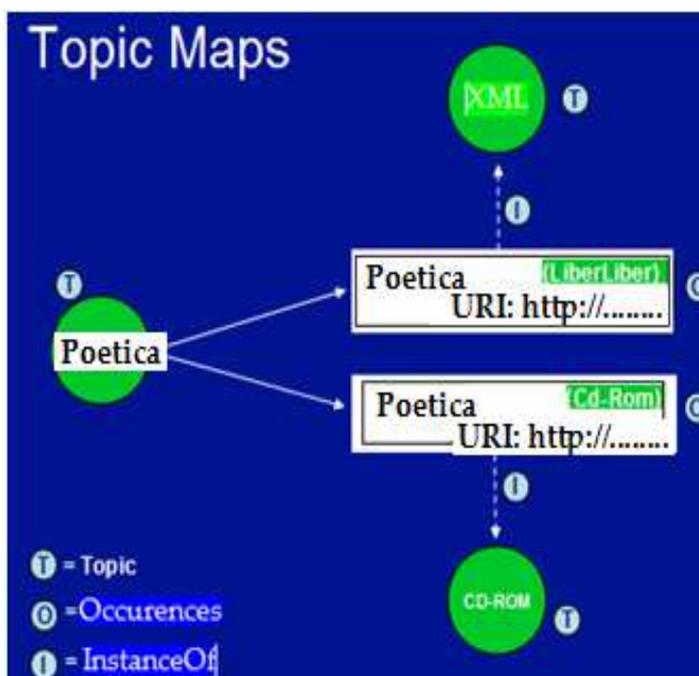


Figura 69. Livello astratto di diversi *Topic*, con Istanze e Associazioni. Occorrenze di *Topic* appoggiati sugli identificatori Web URI

Per ridurre eventuali problemi di ridondanza, è opportuno compiere *operazioni di fusione* tra mappe differenti o porzioni di mappe. Questo consente di evitare di replicare porzioni di mappa che potrebbero essere invece importate dall'esterno, agevolando così anche la riusabilità dei contenuti. È possibile astrarre la mappa dei *Topic* dal contesto delle occorrenze e utilizzarla su strutture di risorse informative differenti, oppure predisporre delle mappe differenti per la stessa struttura di risorse.

Le TM rappresentano una tecnologia opportuna per *esprimere e collegare* sul Web le *descrizioni* provenienti da *diversi domini* come, ad esempio, quelle *archivistiche ISAAR* e bibliografiche (via i metadati e modelli concettuali MARC, FRBR e RDA) [Oh, 2008].

Ad un livello, quindi, la tecnologia TM può permettere l'avvicinamento di più record di metadati e creare collegamenti tra di loro per mezzo di uno strato comune che delinea le Associazioni e le relazioni tra pezzi di contenuto. Ad un altro livello, TM può agire come un'infrastruttura per un *Web front-end* come Portale (es. *National Library of Australia*) [V. sito], raccordando in un unico luogo virtuale diverse risorse digitali distribuiti. Rapportando il paradigma TM al contesto della progettazione sia di contenuti che di servizi a supporto alla didattica, incluso servizi per la formazione a distanza e scenari di ricerca applicata nell'ambito dell'e-learning

potremmo immaginare di avere una Mappa che definisce i *concetti chiave* di un corso o di una lezione e le *relazioni semantiche* tra essi e un *Archivio di materiali didattici* le cui

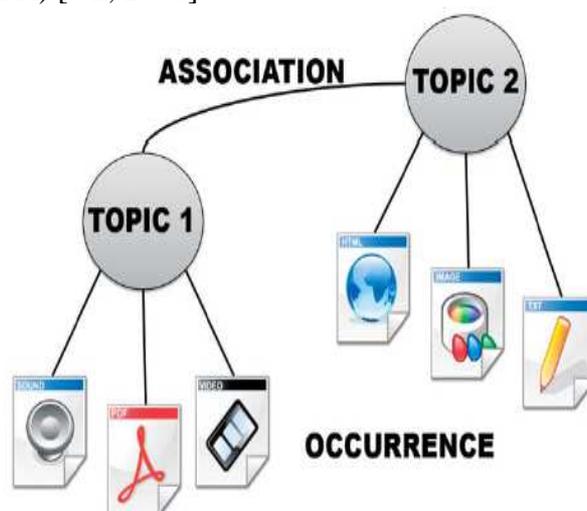


Figura 70. Associazione tra *Topic* su due livelli

risorse sono *associate* ai concetti espressi dalla mappa. L'esistenza dei due livelli indipendenti permette di predisporre approcci conoscitivi differenti e personalizzati alla medesima struttura di risorse didattiche [Adorni et al., 2007, p.54] (Figura 70).

La tecnologia TM presenta uno dei possibili modelli del Web Semantico per la progettazione di alcuni prototipi relativi alla strutturazione e presentazione di descrizioni semantiche, Portali multi-ambito, Sistemi fortemente relazionali e associativi (ontologie) specifici per diverse porzioni di realtà. Le TM possono perciò servire allo scopo della rappresentazione e l'interscambio della conoscenza, offrendo alcuni strumenti per superare problemi tipici nella definizione di Ontologie per l'uso nei sistemi digitali e nella costruzione di strutture che, per loro natura, siano in accrescimento o complesse nell'identificazione, nell'intestazione e nella descrizione dei contenuti digitali.

1.3. *Unified Modeling Language (UML): un modello di classi e loro relazioni per lo sviluppo di Ontologie descrittive*

Nel paragrafo precedente abbiamo visto come le *Topic Maps* possono servire allo scopo di costruzione di *modelli concettuali* rappresentanti un particolare *dominio* ovvero *area di conoscenza* attraverso la definizione di Classi e Sottoclassi, delle associazioni e relazioni che intercorrono tra esse.

Per descrivere domini di diverso tipo possono essere impiegate anche le *teorie logiche* a sostegno della progettazione di oggetti entro i sistemi complessi, per darne una descrizione esauriente. In questa prospettiva, la tecnologia *Unified Modeling Language (UML)* [V. sito] si configura come un potente meccanismo per esprimere *regole logiche* per la rappresentazione di Ontologie e intere Basi di conoscenza (Ontologie + insieme delle istanze delle Classi) per rappresentare i sistemi complessi [Gasevic et al., 2004] (Figura 71).

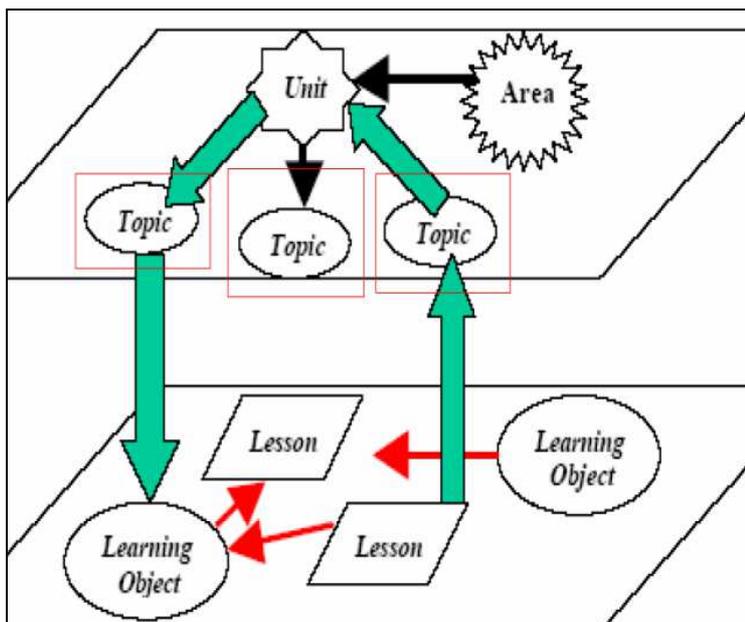


Figura 71. Dai modelli concettuali alle teorie logiche dell' *Unified Modeling Language*

Il linguaggio logico e grafico UML è stato proposto a metà degli anni '90 dall'*Object Management Group (OMG)*, come un codice iconico per la modellazione concettuale di diverse applicazioni (a livello di processi) di Software, basandosi sul paradigma "*Object-Oriented*" (o.o.), che si presta bene per la progettazione e la documentazione di sistemi Software complessi orientati agli oggetti.

UML è stato formalizzato come un *Open Standard ISO/IEC 19501:2005* e si serve della notazione grafica (*diagrammi*) per rappresentare diverse astrazioni del modello del sistema in esame, tramite meccanismi di estendibilità e vincoli espressi in OCL (*Object Constraint Language*) [V. sito].

UML (versione 2.0) prevede l'uso di due gruppi di *diagrammi fondamentali* (con l'uso facoltativo di loro parti costitutive):

I. DIAGRAMMI LOGICI:

- 1) *Use Case* (funzionalità del sistema, analisi dei requisiti del sistema),
- 2) *Class* (struttura statica del sistema: classi del sistema e le relazioni che intercorrono tra loro),
- 3) *Sequence* (interazioni tra elementi del sistema per un obiettivo: i messaggi che i vari elementi si scambiano e in quale ordine),
- 4) *Collaboration* (interazione tra elementi e loro relazioni organizzati nel tempo e nello spazio, per rendere chiare le relazioni più importanti),
- 5) *State Chart* (condizioni di stato e risposte degli elementi del sistema),
- 6) *Activity* (attività degli elementi del sistema nel *Workflow* di processi, con una forte componente di computazione parallela)

II. DIAGRAMMI DI IMPLEMENTAZIONE:

- 1) *Component* (organizzazione degli elementi che realizzano il sistema),
- 2) *Deployment* (configurazione delle componenti fisiche distribuite che realizzano il sistema).

Diagrammi delle Classi UML vengono corredati da descrizioni testuali, combinando le tecniche di:

- modellazione dati (diagrammi “Entità-Relazione” di database);
- modellazione di processi (*Workflows*);
- modellazione di oggetti e loro componenti;
- modellazione logica di domini di conoscenza.

L'insieme di diagrammi UML definisce il modello di applicazione Software.

In Figura 72 viene presentato un possibile impiego dei diagrammi UML, cui Classi rappresentano graficamente alcuni formalismi di *TopicMaps* [V. sito: *Topic Maps — Data Model*].

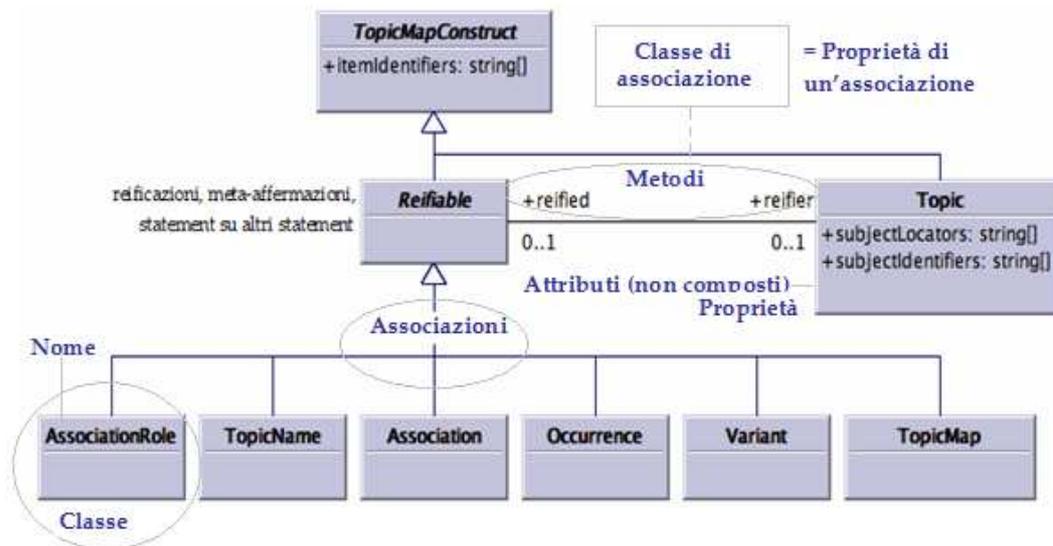


Figura 72. Gerarchia delle Classi di *Topic Maps* rappresentata con il linguaggio UML

In Figura 72, ogni *Classe* viene rappresentata con un rettangolo contenente il *Nome* della Classe (in alto), gli *Attributi* e i *metodi* (es. il metodo *reified* consente di eliminare le relazioni di grado superiore). Le *linee* che collegano due classi rappresentano le *Associazioni*(relazioni) *binarie*. Non si possono assegnare attributi alle associazioni: a tal scopo devono essere usate le *Classi di*

associazione per descrivere proprietà di un'associazione, legate ad essa per mezzo di una linea tratteggiata.

Con gli schemi UML (ordinati in blocchi interagenti) è possibile rappresentare sistemi complessi di diversi domini:

- 1) dalla strutturazione interna dei documenti alla progettazione di schemi di metadati;
- 2) dalla progettazione del sistema "Archivio" alla presentazione di suoi oggetti concreti come, ad esempio, "Risorsa", "Materiale" (Figura 73);
- 3) dalla strutturazione grafica dei formalismi per la rappresentazione della conoscenza come, ad esempio, KIF alla progettazione di architetti Software web, permettendo anche la verifica e generazione automatica di Codice Software.

Considerando tale capacità dell'UML, si potrebbe ipotizzare che i suoi costrutti si possano adeguare facilmente anche alla formalizzazione grafica di tutti i costrutti del Web Semantico (es. *Topic Maps*, RDF, OWL, *Linked Data*), un Web aperto, con i dati interoperabili e dinamici, dove il numero di oggetti digitali cresce a dismisura ed ognuno contiene una molteplicità di istanze (dati, metadati, immagini ecc.) che devono essere raccolte ed elaborate dalle applicazioni *Software*, le quali anche cambiano continuamente e le loro architetture possono diventare, di volta in volta, inconsistenti, inaffidabili o non disponibili.

Le ragioni per cui l'UML può essere considerato una notazione adatta per la modellazione concettuale di Ontologie sono anche essenzialmente legate alla sua relativa immediatezza di utilizzo dei propri costrutti e all'efficienza dei suoi *case-tool* come, ad esempio, *ArgoUML open source UML modeling tool* [V. sito], che è più manovrabile di *ontology-tool Ontolingua* e *Protégé*, i quali richiedono una certa esperienza nel campo della rappresentazione della conoscenza.

Di seguito (Figura 73) vi è un esempio della progettazione tramite UML dell'oggetto del sistema Archivio di ricerca VOA3R "Risorsa":

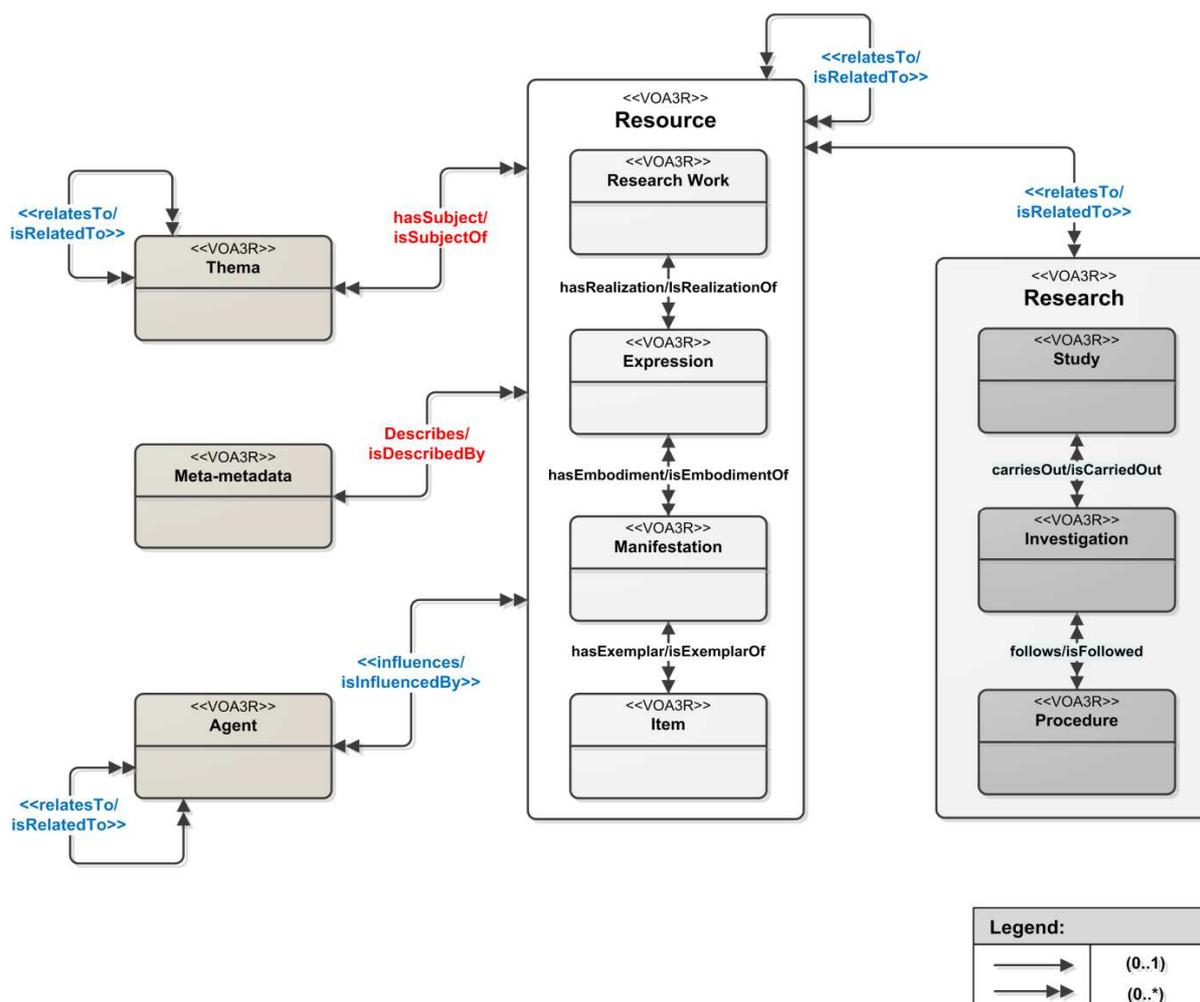


Figura 73. UML per la progettazione dell'oggetto "Risorsa" dell'Archivio di ricerca VOA3R [V. sito : *Specification of metadata profiles and mappings to existing technology*]

In Figura 73 attraverso il modello UML vengono presentati:

1) la Classe "Resource", che fornisce il concetto di base del modello complessivo. Questa entità deve essere in grado di rappresentare una varietà di tipologie di risorse (es. un articolo della rivista): da quelle analogiche a quelle digitali, nonché (meta)dati e relazioni a supporto delle risorse.

La Classe "Risorsa" si esprime attraverso il modello concettuale FRBR (*Functional Requirements for Bibliographical Records*) (APPENDICE B, § 2.2), quindi tramite la realizzazione di una "Expression" che può incorporare una "Manifestation" esemplificata, a sua volta, attraverso un "Item".

2) Ogni "Resource" è correlata alla Classe "Research" che fornisce un segnaposto per la rappresentazione di ulteriori aspetti del ciclo di vita di una ricerca, associati a una o più risorsa/e, protocolli di ricerca, metodi e strumenti utilizzati. La ricerca viene istanziata tramite le entità "Study", "Investigation" e "Procedure".

2) la Classe "Agents" rappresenta individui, gruppi o organizzazioni che influenzano il ciclo di vita di una risorsa: la sua creazione, la pubblicazione, la diffusione, la convalidazione;

3) la Classe "Thema" denota il soggetto, tema, concetto, e la categoria a cui appartiene il contenuto della risorsa;

4) la Classe "Meta-metadata" contiene le informazioni necessarie per la gestione di una collezione (es. creatore, la data di creazione di metadati, annotazioni rilevanti di metadati, ecc.)

Alcune di queste informazioni potrebbero essere anche utili per aiutare gli utenti a stabilire la base per le affermazioni fatte per un record di metadati;

5) le relazioni (*relatedTo*, *hasSubject*, *is Described* ecc.) sono stabilite tra la classe “*Resource*” e le altre classi (e le loro istanze) del modello.

In seguito (Figura 74) verrà dimostrato come il linguaggio grafico UML può essere impiegato per rappresentare l’acquisizione e la catalogazione di materiali e reperti in un Archivio, al fine di programmare i loro collegamenti per architettare meglio una futura ricerca sui contenuti dei medesimi materiali.

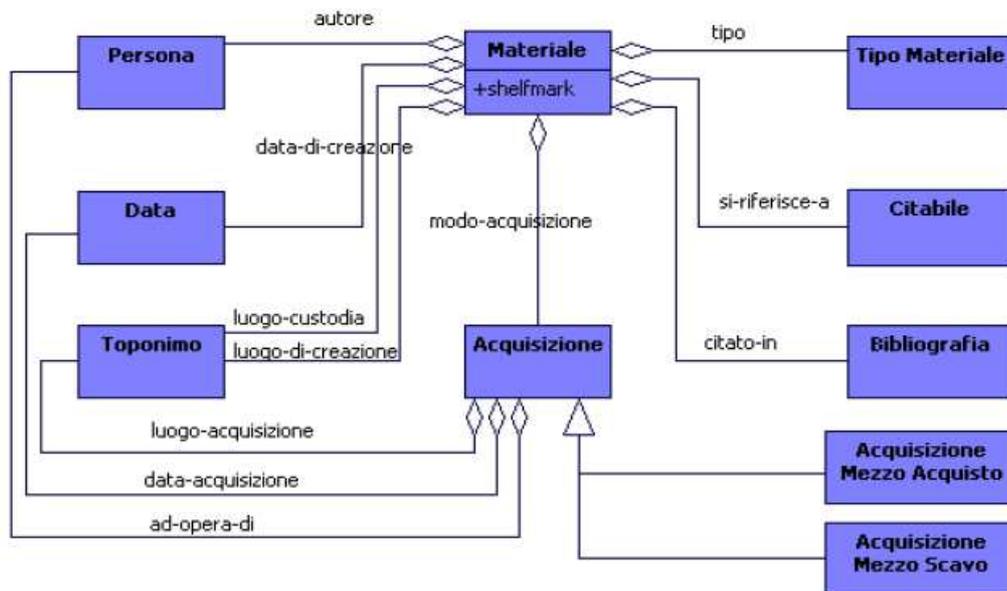


Figura 74. La modellazione dell’oggetto “Materiale” di un Archivio di ricerca

In dettaglio, in Figura 74:

- 1) La classe “MATERIALE” serve a modellare qualunque tipo di oggetto da catalogare, al fine di riuscire ad ordinare materiali e reperti in gruppi omogenei. A questo scopo è necessario associare la classe “Materiale” ad uno o più tipi (Tipo Materiale).
- 2) Affinché sia possibile ricercare i materiali, non solo in base al Tipo, ma anche rispetto ad altre

Caratteristiche, è opportuno introdurre nuove classi e relazioni con cui ad ogni “Materiale” possa venire associato un “AUTORE”, un “LUOGO” e una “DATA” di creazione e un “Luogo” dove viene depositato il medesimo “Materiale”. Inoltre è possibile specificare di cosa parla e da chi viene citato il medesimo “Materiale”.

- 3) La Classe “ACQUISIZIONE” è un’entità a se stante, distinta ma collegata al “Materiale”, che serve a modellare come (Acquisizione Mezzo Scavo, Acquisizione Mezzo Acquisto) uno o più materiali siano stati acquisiti, ad opera di chi, quando ed in che luogo. Qualora si presenti la necessità di rappresentare una diversa modalità di acquisizione, è sempre possibile aggiungere un’ulteriore classe specializzata.

Occorre anche osservare che la modellazione tramite UML è funzionale agli obiettivi del progetto e pertanto può essere rivista, integrata, modificata ed estesa al variare di quest’ultimi.

Uno tra gli strumenti per la modellazione dei modelli di dominio tramite UML è “*Poseidon for UML*” [V. sito] (Figura 75), un *tool* interamente in Java, che lo rende indipendente dalla piattaforma Software e hardware su cui è installato.

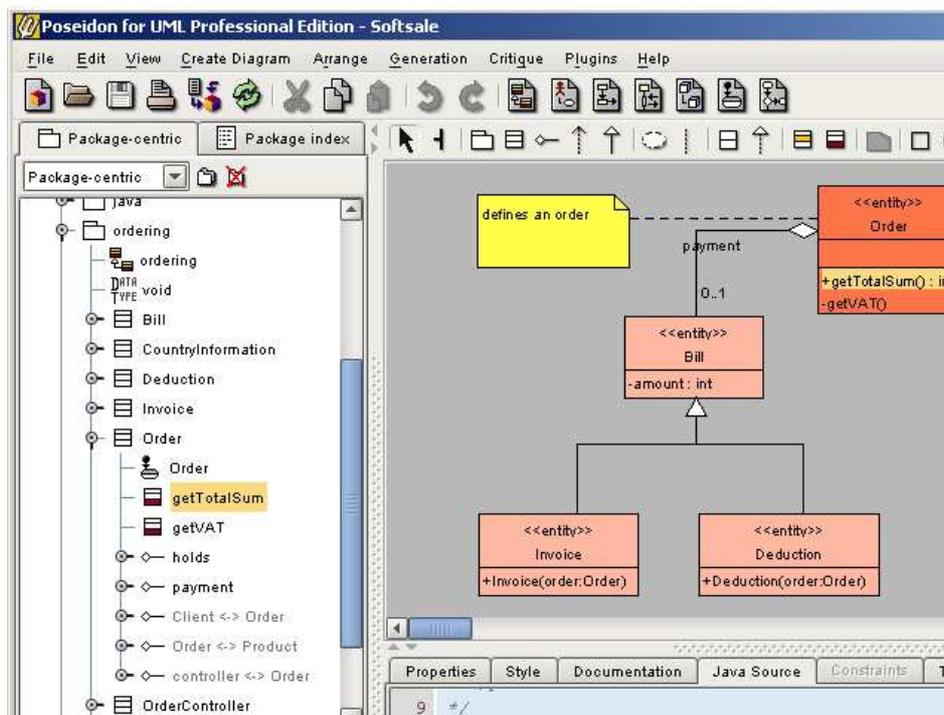


Figura 75. Lo strumento “Poseidon for UML” per la strutturazione grafica “machine-readable” dei modelli di dominio

1.4. Social Web (Web 2.0) per una cooperazione costruttiva dei contenuti digitali

Dalla sua nascita, World Wide Web costantemente consuma, rielabora, produce e condivide informazione in formato digitale. Tali compiti vengono in parte appoggiati sullo spazio della Rete sociale del Web, chiamato Social media o Media Sociali, con i quali si indicano tecnologie e pratiche online che le persone adottano per condividere contenuti testuali, immagini, video e audio.

I Media Sociali sono nati mentre si parlava, si teorizzava, si criticava, si sperimentava, e si costruiva qualcosa di metodologico e tecnico per creare le basi per il *Semantic Web*, e mentre il linguaggio web XML è divenuto uno standard affermato.

I Media Sociali rappresentano un gruppo di applicazioni Internet basate sui presupposti ideologici e tecnologici del Web 2.0 (o *Social Web*) [O'Reilly, 2005] che consentono la creazione e lo scambio di contenuti generati dagli utenti [Kaplan, Haenlein, 2010].

Il Web 2.0 (così come anche *Semantic Web*/Web 3.0) indica uno stato di evoluzione del World Wide Web, rispetto a una condizione precedente. Mentre

il *Semantic Web* nasce dall'alto, direttamente dal W3C; il *Web 2.0* nasce dal basso, da tutta una serie di iniziative eterogenee tra di loro con in comune determinate tecnologie o una particolare visione. Il *Semantic Web* è rappresentato dalle ontologie, un modello entità-relazioni che rappresenta una “specification of conceptualization”, secondo la definizione data nelle scienze dell'informazione; il *Web 2.0* è caratterizzato dalle *folksonomies*, le classificazioni collaborative libere e non gerarchiche basate sui *tag*. Da un lato abbiamo le inferenze basate su procedimenti logici rigorosi che producono nuova conoscenza, dall'altro invece il *Mash-Up*, la combinazione e il riutilizzo di informazioni e servizi [Meschini, 2006, p. 8].

Il Web 2.0 - un “termine-ombrello” come direbbe Umberto Eco, raccoglie e identifica le diverse metamorfosi in atto, in particolare tutte quelle applicazioni online che permettono un elevato livello di interazione tra il sito web e l'utente, come i *Blog*, i *Forum*, le *Chat*, i *Wiki*, le piattaforme di condivisione di media come *Flickr*, *Youtube*, *Vimeo*, i *Social Network* e strumenti

collaborativi come Facebook, Myspace, Twitter, Google+, LinkedIn, Mendeley, ottenute tipicamente attraverso opportune tecniche di programmazione Web afferenti al paradigma del Web dinamico, in contrapposizione al cosiddetto Web statico o Web 1.0.

Esiste qualche punto che accomuna il Web 2.0 e il Web Semantico?

Ad esempio i *feed Rss*, una delle caratteristiche principali dei *Blog*, nascono come applicazione di *RDF*, iniziano a diffondersi i primi *Semantic Wiki*⁵⁰, e all'inizio *Web 2.0* e *Semantic Web* venivano utilizzati come sinonimi [ivi, p. 8].

Il Web 2.0 costituisce anzitutto un approccio filosofico alla rete che ne connota la dimensione sociale, della condivisione, dell'autorialità rispetto alla mera fruizione. Nonostante dal punto di vista tecnologico molti strumenti della rete possano apparire invariati (come Forum, Chat e Blog, che già "preesistevano" nel web 1.0) è proprio la modalità di utilizzo della rete ad aprire nuovi scenari fondati sulla compresenza nell'utente della possibilità di fruire e di creare/modificare i contenuti multimediali.

Attraverso Web 2.0 chiunque è in grado di pubblicare i propri contenuti, dotandoli anche di particolare veste grafica, senza possedere alcuna particolare preparazione tecnica. La tecnologia Wiki (Wikipedia ne è la più illustre applicazione) è il punto di arrivo del *Content Management* (gestione dei contenuti digitali), in quanto ne implementa tutti i paradigmi per la gestione del ciclo di vita dell'informazione per far fruire dell'informazione nell'ambiente stesso in cui essa è nata. In particolare, la possibilità di creazione e condivisione di contenuti su Web, tipica del Web 2.0, è data da una serie di strumenti (*tool*) on-line che permettono di utilizzare il Web come se si trattasse di una normale applicazione.

In aggiunta alla creazione condivisa di contenuti on-line, il Web 2.0 è caratterizzato dalla pubblicazione immediata dei medesimi contenuti e alla loro immediata classificazione e indicizzazione nei motori di ricerca, in modo che l'informazione sia subito disponibile a beneficio dalla comunità, realizzando in maniera veloce il ciclo di vita del *Content Management*. Ignorare l'esistenza e l'impatto di questo fenomeno comunicativo, che ha contribuito abbondantemente alla crescita di un "information age" distribuita e cooperativa, è praticamente impossibile: "some Social Networks have been created and tailored to scientists' needs, in order to make them find researchers with similar interests or expertise, to keep in touch with their peers, to share their information" [Giglia, 2011, p.148].

In vari sistemi di gestione dell'informazione, compreso gli Archivi Istituzionali, vengono implementati o fatti riferimenti a diverse funzionalità del Web 2.0. Negli Archivi Istituzionali, tra le più diffuse implementazioni del Web 2.0, vi sono "Cloud Tags" (le nuvolette di marcatori associati ad una risorsa o un concetto), gli spazi commentare e annotare gli oggetti, creare collezioni personalizzate da utilizzare a fini didattici o di ricerca, gli spazi per la condivisione dei contenuti tramite *Social networks* come *Twitter*, *Facebook*, *Mendeley*, considerati i punti di contatto più rapido e immediato con i diversi lettori, partecipando così allo sviluppo di una ampia comunità scientifica on-line diversificata dalla commistione e dall'ibridazione di vari saperi (V. APPENDICE A, § 1.4).

APPENDICE B

Alcuni approcci a supporto della programmazione e l'interoperabilità dei contenuti digitali di ricerca

2.1. Principi Internazionali di Catalogazione dell' IFLA (*International Federation of Library Associations and Institutions*)

La “Dichiarazione di Principi Internazionali di Catalogazione” [V. sito] – che ha sostituito ed esteso “Principi di Parigi” (approvati nel 1961) – ha lo scopo di servire come base per la standardizzazione internazionale della catalogazione (sviluppo di codici di catalogazione per i cataloghi di vecchia o nuova concezione con le opere testuali), basandosi sulle grandi tradizioni catalografiche del mondo e sul modello concettuale di *Functional Requirements for Bibliographic Records (FRBR)* dell'IFLA (*International Federation of Library Associations and Institutions*).

Sebbene le diversità culturali, di tradizione catalografica, ma anche di tradizione nel trattamento dell'informazione in generale, gli OPAC (*Online Public Access Catalogues*) in tutto il mondo seguono i principi generali della “Dichiarazione” per supportare in modo univoco e a un grado elevato di qualità i propri servizi.

Nonostante gli Archivi aperti del materiale di ricerca seguano pratiche diverse da quelle offerte per le prassi di catalogazione (la meta-datazione verso la catalogazione)

Ai bibliotecari va detto che un ARCHIVIO NON È UN CATALOGO, è uno strumento fondamentalmente diverso e quindi sia l'interfaccia sia i metadati devono essere “a portata di mano” dell'utente finale, così come lo schema disciplinare per il *browsing* utile alla collocazione entro uno schema disciplinare o anche puramente gerarchico che dovrà essere il più semplice possibile [De Robbio, 2007, p.106].

Alcuni principi della “Dichiarazione”, quelli pragmatici che “sistematizzano le prassi in un modo logico e funzionale” [Guerrini, 2011, p. 10] possono applicarsi anche ai contenuti di Archivi Istituzionali, al fine di migliorare la costruzione delle strutture funzionali di mediazione tra questi contenuti e utenti.

1. *Principi generali:*

- a) tenere presente *l'interesse dell'utente* durante la creazione delle descrizioni e delle forme controllate dei nomi quali punti di accesso;
 - b) *significatività* delle descrizioni e delle forme controllate dei nomi;
 - c) le descrizioni e la creazione dei punti di accesso devono contribuire maggiormente all'*economia complessiva* (al costo minore o all'approccio più semplice);
 - d) le descrizioni e la creazione dei punti di accesso devono essere il più possibile *coerenti e normalizzate* (seguire gli *standard*);
 - e) l'*integrazione* di tutti i tipi di contenuti deve essere possibile attraverso le regole comuni per le descrizioni delle entità e per le forme controllate dei nomi;
 - f) le *entità descritte* devono essere rappresentate *accuratamente (fedelmente)*;
 - g) il *lessico descrittivo* e nei *punti di accesso* deve essere armonizzato con quello della maggioranza degli utenti (*uso comune*);
 - h) le descrizioni e le forme controllate dei nomi devono contenere solo gli elementi necessari e indispensabili (*sufficienti*) per identificare in modo univoco un'entità e per soddisfare le necessità informative l'utente.
2. *Recupero*: la possibilità per l'utente di *trovare* le risorse bibliografiche in una raccolta come risultato di una ricerca basata sugli attributi o sulle relazioni delle risorse.
 3. *Navigazione*: la possibilità per l'utente di *navigare* in e al di fuori di una collezione di contenuti (cioè tramite la disposizione logica dei dati bibliografici (metadati) e di autorità e la presentazione di chiare modalità per muoversi, compresa la presentazione delle relazioni tra enti, concetti, oggetti, eventi, persone e luoghi).

2.2. Famiglia di modelli concettuali FRBR (*Functional Requirements for Bibliographical Records*)

Per essere all'unisono con l'universo bibliografico (*“Il regno relativo alle raccolte delle Biblioteche, degli Archivi, dei Musei e delle altre comunità dell'informazione”*) [V. sito: Dichiarazione di Principi Internazionali di Catalogazione, p. 13] è auspicabile che gli Archivi Istituzionali prendano in considerazione non solo la *“Dichiarazione di Principi Internazionali di Catalogazione”*, ma anche alcuni importanti modelli a supporto di metadati descrittivi che rientrano nella *“scienza degli indici”* e, cioè, nel dominio della bibliografia [Serrai, 2009].

Nel momento in cui si è passati dalla scelta di metadati alla loro formalizzazione logica, vi è la necessità di utilizzo di un modello concettuale per realizzare al meglio un programma di gestione degli stessi indici che indicano le informazioni associate a un oggetto (risorsa) informativa, descrivendone contenuti e proprietà.

I requisiti funzionali per record bibliografici (FRBR) rappresentano un modello concettuale che si concentra sulla rappresentazione dell'universo bibliografico, utilizzando un modello *“Entità”-“Relazione”*.

Uno degli aspetti più importanti del FRBR è il fatto che esso non è associato a un particolare Schema di metadati o implementazione.

Recentemente FRBR è stato ampliato con i modelli concettuali:

- *FRAD (Functional Requirements for Authority Data)* [V. sito: Requisiti funzionali per i dati di autorità] per qualificare l'accesso controllato ai descrittori che fungono dai punti di accesso e
- *FRSAD (Functional Requirement for Subject Authority Data)* [V. sito], formando così la cosiddetta famiglia FRBR (Figura 76).

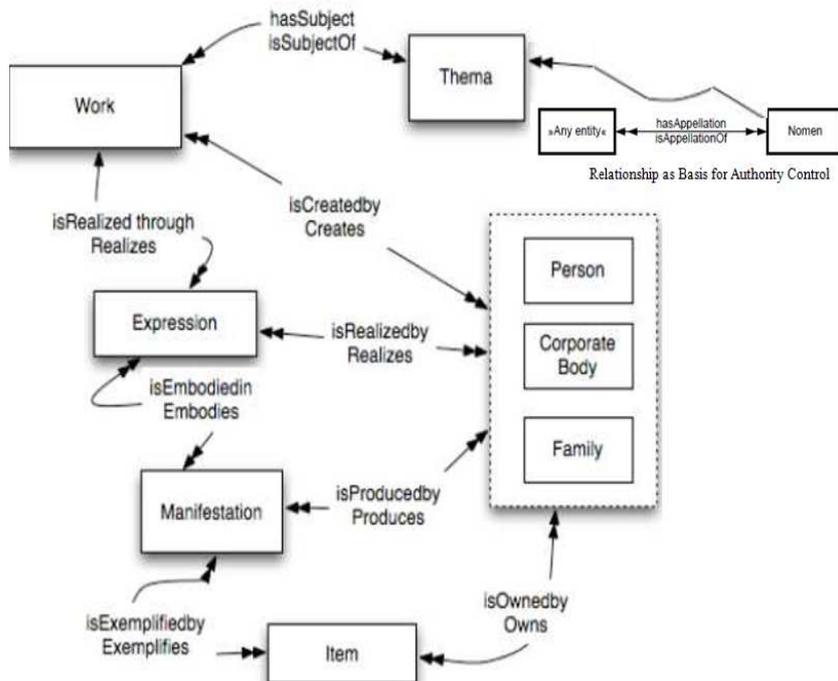


Figura 76. Famiglia dei modelli concettuali FRBR (*Functional Requirements for Bibliographical Records*) [Zumer, Zeng, Salaba, 2010, p.2]

Il modello FRBR è adatto a descrivere tutte le tipologie documentali da quattro diversi ma correlati punti di vista, i quali definiscono un insieme di quattro entità, in riferimento agli aspetti riuniti nei Gruppi sotto i quali può essere concepita una produzione intellettuale.

Gruppo I include:

- OPERA (*Work*): una specifica creazione intellettuale.
- ESPRESSIONE (*Expression*): una realizzazione (*Realization*) di tale Opera.
- MANIFESTAZIONE (*Manifestation*): la materializzazione (*Embodiment*) di un'espressione.
- UNITÀ (*Item*): un singolo esemplare (*Exemplar*) di una manifestazione.

Il Gruppo I viene collegato con il Gruppo II (*Person, Corporate Body*), il quale si riferisce alle entità che possono essere coinvolte nella creazione, realizzazione, distribuzione e gestione delle entità del Gruppo I.

Il Gruppo III si riferisce ai soggetti (*Subjects*) delle opere e comprende: concetti (*concepts*), oggetti (*objects*), eventi (*events*), luoghi (*places*).

FRBR definisce inoltre quattro importanti attività degli utenti in riferimento ad ogni entità o relazione:

- (1) trova (*Find*);
- (2) identifica (*Identify*);
- (3) seleziona (*Select*);
- (4) ottieni (*Obtain*).

Un quinto, non ufficiale compito è navigare o mettere in relazione (*Navigate* o *Relate*).

Il modello FRAD della famiglia FRBR è stato sviluppato per coprire principalmente le entità del Gruppo 2 riferendosi all'informazione di autorità (controllo di autorità delle voci). FRAD ha aggiunto una nuova entità "Famiglia" nel Gruppo 2, al fine di rispondere alle esigenze descrittive di collezioni di archivi. Inoltre, FRAD ha definito una serie di nuove entità, attributi e relazioni che corrispondono alle esigenze di dati di autorità, come *Nome* con cui un'entità (un'Opera) è conosciuta nel mondo reale, *Identificatore* (un numero, codice, una frase, univocamente associate ad una entità), *Punto di accesso controllato* (il termine con il quale una risorsa può essere trovata), *Regole* (le istruzioni per la formazione di un punto di accesso controllato), e *Agenzia* (un'organizzazione che crea e mantiene i punti di accesso controllati).

Il modello concettuale FRSAD è un'altra estensione di FRBR sviluppato in parallelo a FRAD. Il modello è focalizzato a identificare come *entità, attributi e relazioni* sono correlati a *dati di autorità* riferiti al *Soggetto* (tema) di una risorsa. FRSAD ha definito due entità: *Tema* (qualsiasi entità usata come l'oggetto di un'Opera) e *Nomen* (qualsiasi segno o una sequenza di segni attraverso cui un Tema è conosciuto, riferito o affrontato). L'inclusione di queste entità è necessaria per modellare il processo di controllo di autorità, un importante meccanismo di controllo di varianti delle denominazioni. Se il controllo di autorità non viene implementato, la denominazione (titolo, nome, soggetto, ecc) di qualsiasi risorsa può essere modellata come un attributo di tale questa risorsa, invece di essere un'entità a sé stante. Tale pratica non solo limita la possibilità di definire gli attributi di denominazione e le relazioni tra le diverse denominazioni, ma mescola anche le relazioni semantiche di inter-concetto con le relazioni lessicali di inter-denominazione.

2.3. Dublin Core Application Profile (DCAP): creazione di profili di metadati applicativi per domini specifici

Il *Dublin Core Application Profile* (DCAP) [V. sito] è un profilo applicativo proposto entro le iniziative della comunità *Dublin Core Metadata Initiative* (DCMI), attraverso i documenti come "Singapore Framework for Dublin Core Application Profiles" e "Guidelines for Dublin Core Application Profiles".

Il “*Singapore Framework*” [V. sito] definisce un insieme di componenti descrittivi necessari o utili per la documentazione di un profilo di metadati applicativo, circoscrivendo come questi elementi si relazionano con i modelli standard di un dominio e con i fondamentali Standard del Web Semantico (RDF, OWL, SKOS).

Il documento “*Guidelines for Dublin Core Application Profiles*” [V. sito] definisce la strutturazione dei processi di un profilo applicativo DCAP, attraverso cui è possibile descrivere un qualsiasi dominio di contenuti per una particolare applicazione (Figura 77):

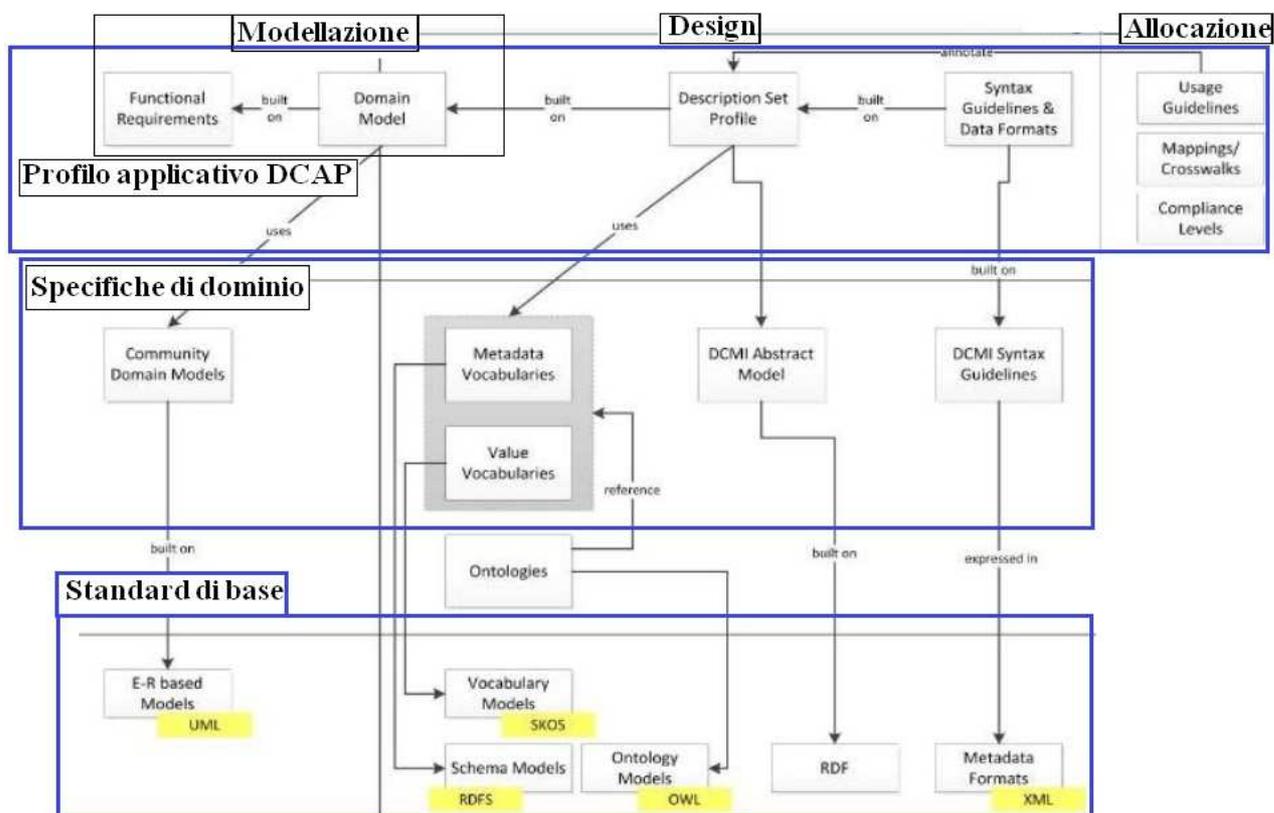


Figura 77. Modello DCAP (*Dublin Core Application Profile*) per la modellazione di profili applicativi di metadati a supporto di diverse tipologie di contenuti

Nella Figura 77, la sezione “Allocazione” comprende:

- 1) “*Usage Guidelines*” (enumerazione di termini di metadati da utilizzare e le regole per il loro impiego in un contesto applicativo);
- 2) “*Mappings/Crosswalks*” (mappatura tra metadati);
- 3) “*Compliance Level*” (compilazione di metadati).

La sezione “Design” comprende le sezioni:

- 1) “*Description Set Profile*” (un metodo per specificare vincoli strutturali sulle descrizioni e dichiarazioni ricavate da un record di metadati);
- 2) “*Syntax Guidelines and Data Formats*” (la sintassi che sarà utilizzata dalla macchina per codificare i dati).

La sezione di “Modellazione” include:

- 1) il modello di dominio semplice (“*Domain Model*”), per caratterizzare le tipologie di oggetti che devono essere descritte da metadati e le loro relazioni e
- 2) un modello di “*Functional Requirements*”- basato su FRBR - per definire più in dettaglio quale modello di contenuti sarà realizzato da una comunità.

Ogni sezione, a sua volta, si suddivide in altre sottosezioni raggruppate sotto i livelli “Specifiche di dominio” e “Standard di base”, con l’applicazione si può modellare qualsiasi profilo di metadati.

Tra i modelli applicativi sviluppati sulla base di DCAP, ci sono:

- *Dublin Core Collections Application Profile* (DC-PAC), per la costruzione di un set di metadati a supporto della descrizione di livelli di una collezione;
- *Images Application Profile* (IAP), per il supporto della descrizione di immagini in Archivi istituzionali;
- *Scholarly Works Application Profile* (SWAP), per il supporto della descrizione di contenuti di ricerca in Archivi istituzionali;
- *Dublin Core - Library Application Profile* (DC-Lib), per il supporto della descrizione di oggetti in applicazioni correlate con il dominio di biblioteche;
- *Dublin Core -Education Application Profile* (DC-Ed AP), per il supporto della descrizione di oggetti didattici.

2.3.1. Profilo applicativo per contenuti di ricerca *Scholarly Works Application Profile* (SWAP)

Il Documento *Scholarly Works Application Profile* (SWAP) [V. sito] è stato definito da *Dublin Core Metadata Initiative* nel 2009 come un'applicazione pratica di *DCAP* per descrivere lavori scientifici (articoli di rivista *peer-reviewed*, *pre-print*, *working paper*, tesi, parte del libro, rapporto scientifico ecc.), riconoscendo tali come una distinta creazione intellettuale o artistica accademica prodotti da studiosi e ricercatori.

SWAP si appoggia sul già descritto modello FRBR: 'Work', 'Expression', 'Manifestation' e 'Item', e include inoltre la quinta Entità 'Agent' ('Person' e 'Corporate Body' in FRBR) (Figura 78).

Le descrizioni realizzate con il profilo SWAP sono costruiti principalmente per:

- fornire un profilo di metadati più ricco, consistente ed estensibile del *Dublin Core simple* per la descrizione di lavori scientifici in formato digitale;
- facilitare la ricerca, navigazione o selezione dei contenuti di ricerca;
- supportare l'uso di vocabolari controllati e di controllo di autorità di metadati (*Authority Control*);
- consentire l'identificazione di più recenti versioni (di coppie più appropriate) di risorse;
- sostenere servizi a valore aggiunto, in particolare quelli basati sull'uso di *OpenURL ContextObjects* [V. sito: ANSI/NISO Z39.88 -2004];
- implementare un metodo univoco/non ambiguo per identificare *full-text* di risorse digitali;
- consentire l'identificazione della provenienza, del finanziatore della ricerca e del codice del progetto;
- facilitare l'identificazione dei materiali ad Accesso Aperto (*access and rights management*).

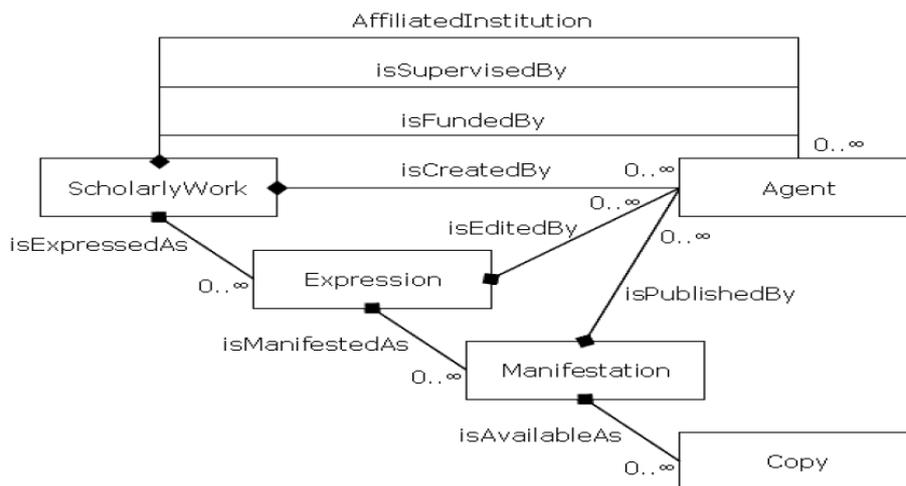


Figura 78. Profilo Applicativo del Modello di dominio SWAP (*Scholarly Works Application Profile*) [Allison, Johnston, Powell, 2007]

Il modello SWAP è stato esteso da Žumer et al. [2010] – in adeguamento al modello della “FRBR Family” - come segue (Figura 79):

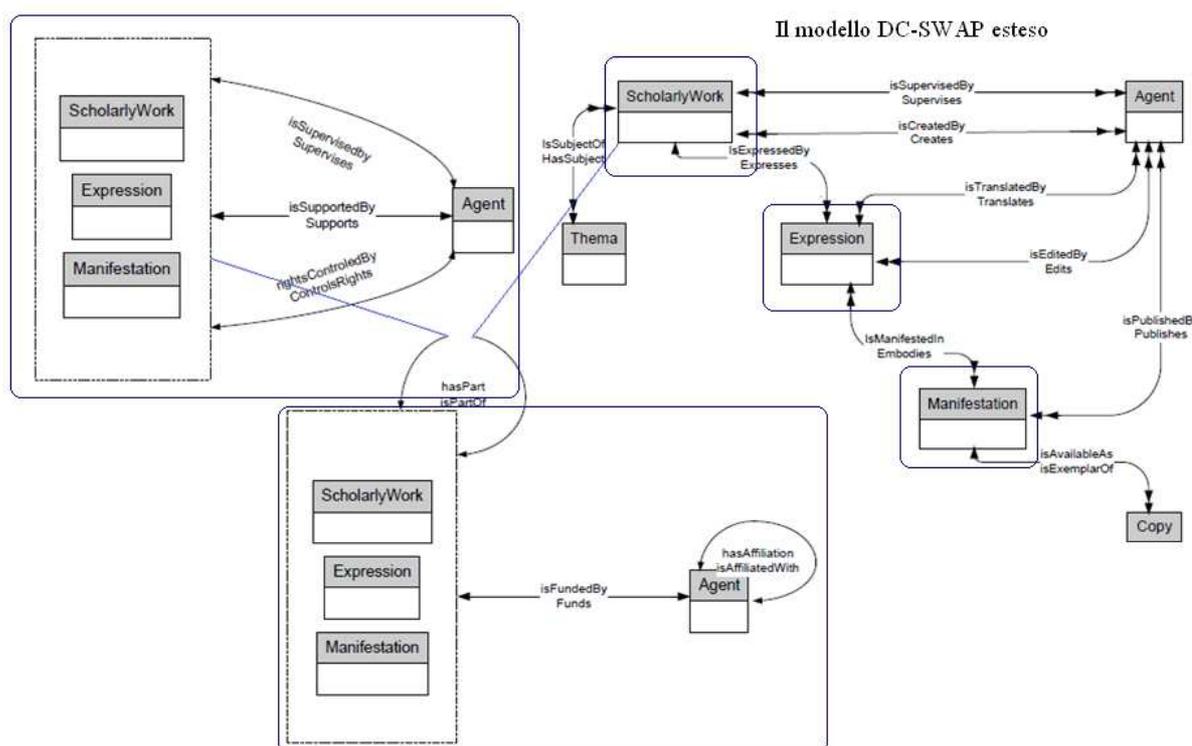


Figura 79. Il modello DC-SWAP esteso [Žumer, Zeng, Salaba, 2010]

Il cambiamento più importante è la costituzione della relazione tra le entità *Expression* e *Manifestation*; oltre al fatto evidente che qualsiasi istanza di *Expression* può essere radicata in ("isManifestedIn") una o più *Manifestation*, e ogni istanza della *Manifestation* può incarnare uno o più istanze di *Expression*.

Esempi di entità “*Manifestation*” che includono più entità “*Expressions*” sono: un testo con le illustrazioni, un numero della rivista, o qualsiasi raccolta di articoli. La relazione “*isManifestedAs*” è stata cambiata a “*isManifestedIn*”.

Viene inoltre aggiunta la relazione “*hasSubject*” tra le entità “*Work*” e “*Thema*”. Le relazioni “*HasPart*” e “*isFundedBy*” possono essere entrambe applicate a qualsiasi entità del GRUPPO 1: “*ScholarlyWork*”, “*Expression*”, “*Manifestation*”.

Le relazioni tra GRUPPO 1 e l’entità “*Agent*” possono essere ulteriormente generalizzate: “*isSupportedBy*” può includere un finanziamento o un supporto non-materiale, la relazione “*isSupervisedBy*” può includere una supervisione formale (es. consulenza di tesi) o informale; relazione “*rightsControlledBy*” serve come meccanismo per l’attribuzione di diritti. Le implementazioni delle relazioni generali illustrati nella Figura, possono essere sostituite da quelle più specifici.

2.4. Schema di metadati proposto dal Gruppo *Open Access* (CRUI, Italia, 2012)

Le “Linee Guida per la creazione e la gestione di metadati nei *Repository* Istituzionali” [V. sito] hanno l’obiettivo di suggerire agli Atenei italiani uno schema di metadati di base il più possibile interoperabile nell’ambito della gestione degli Archivi Istituzionali. Tale schema comprende 33 termini di metadati (descrittivi e strutturali) e quelli legati al contesto di appartenenza (es. informazioni amministrativo-gestionali relative ad afferenza dipartimentale, Settore Scientifico Disciplinare/SSD), affrontando inoltre le problematiche legate al controllo dei nomi (*Authority Control*), del *versioning* (versioni di contenuti depositati) e degli identificatori per ciascuna entità coinvolta nel processo di archiviazione dei dati.

Il riferimento allo schema di metadati delle “Linee” è stato fatto nel Capitolo 5 (Parte III), al fine di dimostrare come sia possibile allineare tale Schema con le *strategie di codifica* delle proprietà di metadati e del loro potenziamento semantico proposte dalle Raccomandazioni “*Linked Open Data enabled – Bibliographical Data*” (LODE-BD, 2012) [V. sito], per migliorare l’espressività delle risorse digitali.

In seguito (Tabella 20) vengono forniti i termini di metadati e le loro espressioni di codifica, secondo lo schema di metadati *Dublin Core* e alcuni metadati PRISM (*Publishing Requirements for Industry Standard Metada*), così come proposto dalle “Linee guida”:

	Termini di metadati	Espressioni di codifica
1.	Abstract	dc:description.abstract
2.	Altri titoli	dc:title.alternative
3.	Autore	dc:creator
4.	Citazione bibliografica	dc:identifier.bibliographiccitation
5.	Contributori	dc:contributor.xxx
6.	Curatori	dc:contributor.editor
7.	Data Congresso	dc:description.conferenceDate
8.	Data di accettazione	dc:dateAccepted
9.	Data di deposito	dc:dateSubmitted
10.	Data pubblicazione	dc:date.Issued
11.	Diritti	dc:rights
12.	Editore	dc:publisher
13.	Edizione	prism:edition
14.	Embargo	dc:date.embargoEnd
15.	Fonte	dc:source
16.	Formato	dc:format

17.	Identificatore Progetto	dc:relation
18.	Lingua	dc:language
19.	Localizzazioni/riferimenti alla risorsa	prism:doi; prism:url; prism:issn; prism:eissn
20.	Luogo Congresso	dc:relation.conferencePlace
21.	Numero del fascicolo	prism:number
22.	Pagina finale	prism:endingpage
23.	Pagina iniziale	prism:startingpage
24.	Presenza e accessibilità del full-text	dc:accessrights
25.	Pubblicazione	dc:relation.ispartof
26.	Relatore	dc:contributor.advisor
27.	Soggetti	dc:subject; dc:subject.ssd; dc:subject.isicrui; dc:subject.ddc
28.	Stato della pubblicazione	dc:type
29.	Tipo di Pubblicazione	dc:type
30.	Titolo del Congresso	dc:relation
31.	Titolo proprio	dc:title
32.	Verificato da Referee	dc:type.refereed
33.	Volume	prism:volume

Tabella 20. Termini di metadati proposte dalle “Linee Guida per la creazione e la gestione di metadati nei *Repository* Istituzionali”

2.5. Core Scientific Meta-Data Model (CSMD)

Il modello *Core Scientific Meta-Data (CSMD)* [V. sito] è un modello orientato allo studio di dati, indirizzato ad acquisire le informazioni di alto livello sulla ricerca scientifica e di dati che essa produce. Le entità del modello (Figura 80) sono a supporto di dati raccolti all'interno delle strutture scientifiche di grandi dimensioni.

Il CSMD viene applicato nelle "scienze strutturali" (es. la chimica, scienza dei materiali, scienze della terra, e biochimica) che si occupano dei temi come la struttura molecolare delle sostanze, all'interno delle quali vengono effettuate le analisi sperimentali sistematiche su campioni di materiali.

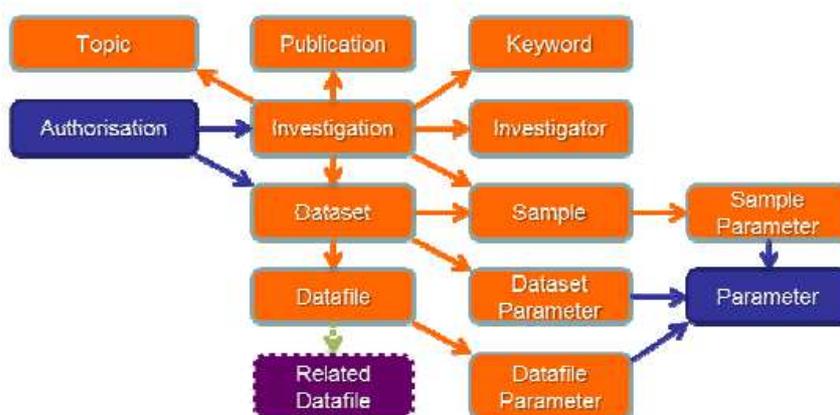


Figura 80. Le entità del modello CSMD per l'acquisizione delle informazioni di alto livello su studi scientifici

Per descrivere il modello a diversi livelli di astrazione viene utilizzato il *DCMI Singapore Framework* (V. pres. Cap., § 2.3) che fornisce una ripartizione delle fasi necessarie per documentare un qualsiasi modello di metadati basato sul DCAP (*Dublin Core Application Profile*).

2.6. Interoperabilità, integrazione, condivisione di contenuti di ricerca

L'Agenda Digitale Europea [V. sito] - una tra sette iniziative presentata dalla Commissione europea per la strategia "Europa 2020" - fissa obiettivi per la crescita nell'Unione Europea (UE). Secondo quanto previsto dall'Agenda, ogni paese membro dell'UE deve allineare entro il 2013 il proprio modello (*framework*) di interoperabilità nazionale al modello europeo di interoperabilità (EIF: l'*European Interoperability Framework*) [V. sito], il quale definisce i livelli di interoperabilità organizzativa, tecnica e semantica.

L'interoperabilità è la possibilità, offerta alle istituzioni, di elaborare informazioni da fonti esterne o secondarie senza perdere il significato delle informazioni stesse nel processo di elaborazione. Questa definizione - offerta dall'Agenda Digitale Europea - assume il ruolo centrale per la collaborazione e l'interscambio di informazioni tra istituzioni pubbliche (amministrazioni) e, quindi, può riguardare anche l'interoperabilità tra informazioni di Archivi Istituzionali.

D'altra parte, gli Archivi Istituzionali *Open Access* hanno già intrapreso la strada di disseminare i loro contenuti tramite il protocollo dell'interoperabilità tecnica *Open Archive Initiative Protocol for Metadata Harvesting* (OAI-PMH) (V. Parte I, Cap.2, § 2.1).

Inoltre, entro diversi Archivi vengono adottati i vari modelli di gestione dati che favoriscono l'interoperabilità semantica tra le informazioni entro e tra gli Archivi. Affinché i modelli adottati possano assicurare la stabilità dell'interoperabilità nel tempo, essi devono essere ben formalizzati entro i requisiti di progettazione di metadati a supporto di contenuti di ricerca, per fornire un quadro di riferimento completo per la produzione di contenuti digitali interoperabili.

La condivisione dei dati della ricerca costituisce oggi una nuova sfida per la scienza che può beneficiare dalle strategie di interoperabilità tra i sistemi, per risolvere problemi complessi di interazione tra le informazioni.

Presupposti per affrontare questa sfida a lungo termine sono:

- 1) il supporto tecnico per una gestione uniforme (standard) dei dati che favorisca il riuso delle informazioni;
- 2) lo sviluppo di infrastrutture che assicurino la conservazione dei dati a lungo termine;
- 3) la predisposizione di politiche idonee all'Accesso Aperto dei dati e al riconoscimento del loro valore per l'interoperabilità semantica che può essere completa solo attraverso la connessione e comunicazione di dati/contenuti aperti.

Nei paragrafi successivi vengono:

- ✓ presentati alcuni scenari attuali che provengono dalle iniziative nazionali e internazionali a supporto dell'interoperabilità semantica e di condivisione dei dati che possono essere adottati in Archivi Istituzionali;
- ✓ descritti i requisiti di interoperabilità semantica che possono far parte della progettazione di metadati a supporto dei contenuti di ricerca;
- ✓ presentato il caso dell'Archivio di ricerca VOA3R in cui i meccanismi di interoperabilità vengono curati attraverso la costruzione dello scenario sociale (alcune funzionalità di Web 2.0 – *Social Web*) e quello semantico (l'implementazione dei meccanismi *Linked Data* – una presentazione pratica del Web Semantico, Web 3.0), tramite cui ai contenuti dell'Archivio viene garantita l'interoperabilità semantica, così come definita dall'*European Interoperability Framework* (EIF).

2.6.1. Rapporto “*The Current State of Open Access Repository Interoperability*”

Tutti i Repository *Open Access*, da quelli Istituzionali a quelli Disciplinari (nazionali o multinazionali) devono supportare l'interoperabilità tra diversi sistemi a loro supporto su di una medesima rete, al fine i loro contenuti possono far parte dell'aggregato globale della conoscenza *Open Access*.

La costruzione dell'infrastruttura di interoperabilità semantica tra i contenuti di *Repository* è un compito ancora relativamente nuovo e, a prima vista (dallo stato dell'arte fornito dal Rapporto “*The Current State of Open Access Repository Interoperability*”), può apparire anche caotica, confusa e complessa.

Il menzionato Rapporto è un elaborato della Confederazione di Repository *Open Access* (COAR) [V. sito: *The Current State of Open Access Repository Interoperability*] del 2012 che fornisce una visione d'insieme dei servizi per l'interoperabilità semantica implementati in diversi Archivi *Open Access*.

In particolare, il Rapporto copre sette aree di interesse per le iniziative di interoperabilità in corso, e fornisce una panoramica di diciannove importanti azioni di interoperabilità, identificando le sfide principali e le aree emergenti nel campo di ricerca sull'interoperabilità semantica delle informazioni.

Il Rapporto indica quali iniziative di interoperabilità sia necessario attuare per i *Repository* al fine di creare i servizi specifici di interoperabilità, offrendo poi attraverso essi servizi di ricerca dell'informazione a valore aggiunto agli utenti.

Le aree che contribuiscono alla qualità dell'interoperabilità tra i contenuti di Repository *Open Access* - coperte dal Rapporto - sono seguenti:

- 1) la raccolta/aggregazione di metadati (*Metadata harvesting*);
- 2) le statistiche d'uso (*Usage statistics*);
- 3) il trasferimento incrociato di contenuti tra i sistemi di Repository (*Cross system content transfer*);
- 4) la creazione di identificatori unici per gli autori e contenuti Repository (*Creating unique identifiers for Authors and Repository content*);
- 5) la necessità di aumentare la visibilità dei contenuti depositati in Repository attraverso i motori di ricerca Internet (*Increasing visibility of contents stored in Repositories through Internet search engines*);
- 6) i servizi di supporto per le reti di Repository (*Support services for Repository networks*);
- 7) l'integrazione di Repository nei flussi di lavoro di gestione di ricerca (*Integrating Repositories into research administration workflows*);
- 8) il supporto per gli oggetti composti - oggetti digitali che includono più componenti o file (*Support for compound objects – digital objects that include multiple components or files*).

Un'ulteriore area attualmente in fase di sviluppo e brevemente affrontata dal Rapporto riguarda i servizi a supporto di *Linked Data* [V. APPENDICE A].

Il Rapporto presta attenzione alla descrizione di sette importanti iniziative che contribuiscono all'interoperabilità tra contenuti digitali di diversi Repository (Figura 81):

- 1) la raccolta/aggregazione di metadati (il protocollo di interoperabilità tecnica *OAI-PMH*);
- 2) le reti di Repository (*DRIVER, OpenAIRE, National repository networks, UK RepositoryNet+*);
- 3) le statistiche d'uso di metadati (*COUNTER, KE-Usage Statistics, OA-Statistik PIRUS/IRUS-UK, SURE*);

- 4) il trasferimento incrociato di contenuti tra i sistemi di Repository (il protocollo *SWORD*, *Open Access Repository Junction*, *CRIS-OAR Metadata Interoperability*);
- 5) gli identificatori univoci per gli Autori (i servizi *AuthorClaim*, *ORCID*);
- 6) gli identificatori persistenti per i contenuti (*DOI*, *PersID*, *Handle System*, *DataCite*);
- 7) la gestione di oggetti composti (il protocollo *OAI-ORE*).

Il *Rapporto* è una delle diverse azioni informative volte a promuovere la comprensione dalla comunità di Repository dello stato dell'arte dell'interoperabilità tra i contenuti digitali, a condividere le migliori pratiche attuali relative all'interoperabilità, e a fornire le conoscenze e le informazioni necessarie per progettare e attuare strategie istituzionali per l'interoperabilità.

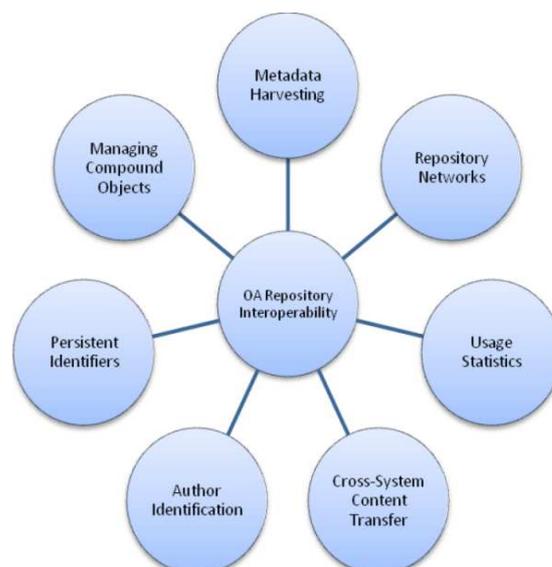


Figura 81. Iniziative della Confederazione di Repository *Open Access* per l'interoperabilità tra contenuti digitali di Repository

2.6.2. Standard *Common European Research Information Format* (CERIF) a supporto dell'interoperabilità tra vari sistemi informativi di ricerca

Per migliorare e rafforzare la creazione e la condivisione dei contenuti di ricerca, a livello di ogni singola nazione esiste un CRIS (*Current Research Information System*), ossia un canale informativo dedicato ad offrire accesso alle informazioni che provengono dai progetti di Ricerca e Sviluppo (R&S). Un CRIS è formato da un *modello di dati* che descrive i vari progetti in via di sviluppo o già realizzati e da un *tool* (o un insieme di *tool*) per maneggiare questi dati, con lo scopo di:

- 1) fornire una "interfaccia comune" (uniforme, almeno a livello riassuntivo) per gestire, integrare e comparare le informazioni scientifiche derivanti dalle diverse fonti di dati;
- 2) migliorare la comunicazione fra scienziati riguardo a progetti in fase di realizzazione;
- 3) offrire una base di informazioni agli amministratori di programmi nazionali di ricerca e sviluppo.

Per raggiungere questi obiettivi è stato proposto un formato comune per la gestione delle informazioni di ricerca per tutti gli stati Europei CERIF (*Common European Research Information Format*) [V. sito] (Figura 82.).

CERIF è un modello formale per l'interoperabilità (gli scambi) di informazioni tra vari sistemi informativi di ricerca (CRIS), ed è organizzato intorno a tre principali entità (in verde):

- 1) “PERSON”(persone, con le relative competenze);
- 2) “PROJECT”;
- 3) “ORGANIZATION UNIT” (organizzazioni, con informazioni sulle attrezzature di cui dispongono e sulle attività di cui si occupano);
- 4) e tre entità correlati (in arancione): “RESULT PUBLICATION”, “RESULT PATENT”, “RESULT PRODUCT” (prodotti, brevetti, pubblicazioni e gli altri “risultati” dei progetti di ricerca).

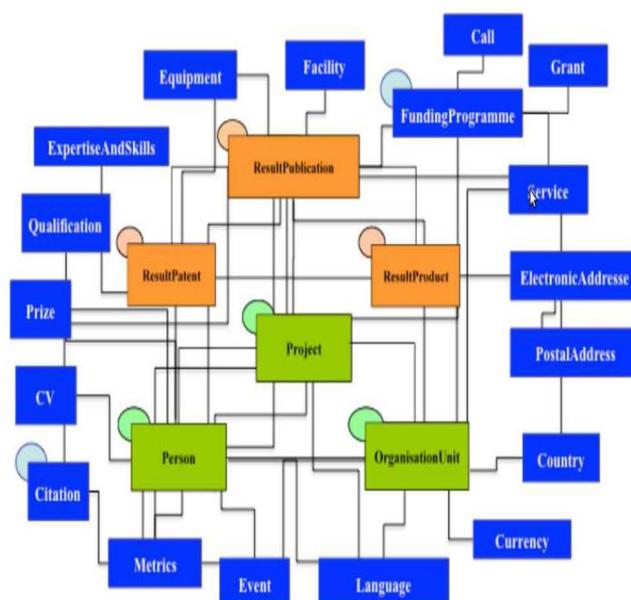


Figura 82. CERIF (*Common European Research Information Format*)

CERIF è un modello multicentrico che si focalizza non solo sulle informazioni che riguardano i dati di ricerca, ma anche sui processi di ricerca a partire dal loro punto di creazione attraverso i processi coinvolti in collaborazione, finanziamento, pubblicazione e comunicazione dei risultati di ricerca. Tale modello attribuisce ai sistemi di informazioni di ricerca le funzionalità sociali, attraverso la gestione delle informazioni che riguardano persone, organizzazioni, progetti di ricerca.

2.6.3. VIVO Ontology: la rappresentazione di entità collaborative coinvolte nei processi di ricerca

L'ontologia VIVO [V. sito] (Figura 83) è un'applicazione *Open Source* per l'attuazione di principi e tecnologie del Web Semantico nella presentazione del contesto di ricerca accademica. In particolar modo, l'ontologia VIVO la quale, attraverso specifiche univoche, formali ed esplicite informazioni descrittive ricercatori, comunità di ricerca accademica, attività e prodotti di ricerca nonché le relazioni che li legano insieme. L'ontologia VIVO permette attività di ricerca trasversali alle diverse discipline, insieme a finanziamenti (compreso la disponibilità di borse di studio), sia in una singola che più istituzioni. Ogni persona, organizzazione o altra entità di dati in VIVO ha un identificatore univoco nella forma di un indirizzo web uniforme URI (*Uniform Resource Identifier*).

Risalendo ad un certo URI nel browser, all'utente verrà restituita una pagina web con i contenuti leggibili, mentre una richiesta per lo stesso identificatore URI come per un insieme di dati da parte di una macchina restituirà una versione di contenuti “*machine-readable*” che potranno essere usati in diverse applicazioni, seguendo i principi di *Linked Open Data* [V. APPENDICE A].

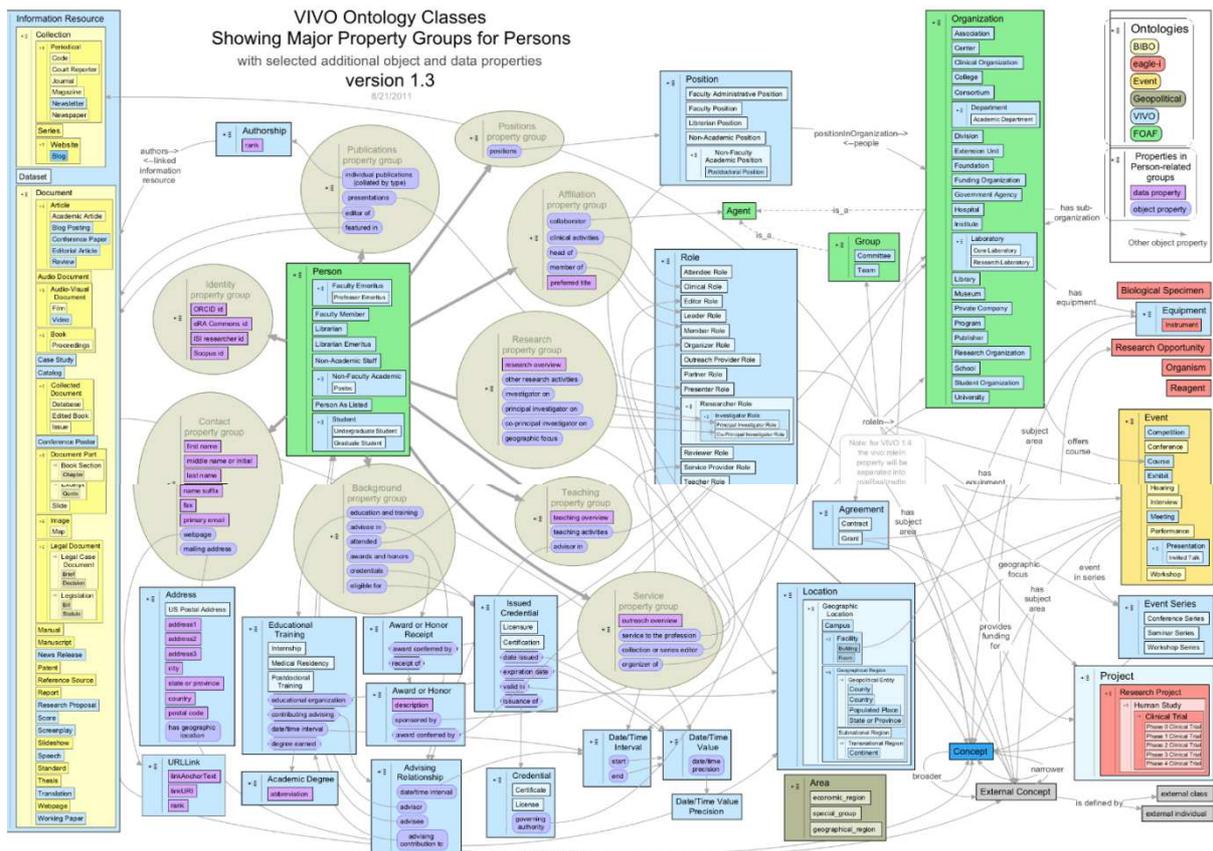


Figura 83. Ontologia VIVO: rappresentazione dei ricercatori e dell'intero contesto in cui essi lavorano

L'ontologia VIVO è strutturata da un insieme di tipi (Classi) e le relazioni (Proprietà) per rappresentare i ricercatori e l'intero contesto in cui essi lavorano. La popolazione e l'aggiornamento di contenuti in una qualsiasi installazione VIVO locale possono essere mantenuti manualmente, nonché in modo automatico portando in VIVO le informazioni da:

- sistemi locali di registrazione, come informazioni su risorse umane, sovvenzioni e borse di studio, corsi, le banche dati di attività di facoltà/dipartimenti, o da
- fornitori di database come aggregatori di pubblicazioni e le agenzie di finanziamento.

L'ontologia VIVO si appoggia su una serie di ontologie di base come *BIBO* (*Bibliographic Ontology*), *FOAF* (*Friend Of A Friend Ontology*), termini di metadati standard *Dublin Core* [V. siti] e altri.

Un buon esempio dell'implementazione dell'ontologia VIVO è presentato dal Portale di ricerca "AgriVIVO" [V. sito] costruito per sviluppare le connessioni tra diversi agenti attivi nel campo di ricerca sull'agricoltura, connettendo gli agenti registrati nei directory separati e in diverse comunità online.

2.6.4. Citation Typing Ontology (CITO)

L'Ontologia "Citation Typing Ontology" (*CITO*) è un linguaggio formale per specificare come una risorsa cita un'altra. Essa contiene diversi termini particolarmente rilevanti per la citazione dei dati, mentre i termini aggiuntivi ("*Other Reasons*") possono essere trovati nell'ontologia estensionale *CiTO4Data* (Figura 84).

13. Peters B, Sidney J, Bourne P, Bui HH, Buus S, Doh G, Fleri W, Kronenberg M, Kubo R, Lund O, The immune epitope database and analysis resource: from vision to blueprint. PLoS Biol. 2005;3:e91. [PMC free article] [PubMed]

Why does this article cite that reference? (Choose as many reasons as apply by clicking on them.)

cites as authority	critiques	obtains background from	uses data from
cites for information	discusses	reviews	The citing entity updates statements, hypotheses or understanding presented by the cited entity.
corrects	extends	updates	
Other Reasons			
agrees with	compiles	documents	refutes
cites as data source	confirms	gives background to	replies to
cites as evidence	contains assertion from	gives support to	retracts
cites as metadata document	credits	has reply	ridicules
cites as recommended reading	derides	includes excerpt from	shares author
cites as related	disagrees with	includes quotation from	supports
cites as source document	disputes	qualifies	uses conclusion

Figura 84. CITO (Citation Typing Ontology): espressione delle citazioni correlate relative a un articolo di ricerca

La CITO utilizza le asserzioni per descrivere le relazioni che esistono tra diversi set di dati, come le opere scientifiche (es. articoli) e le loro citazioni e viceversa, attraverso la formalizzazione di *dataset* e di affermazioni su di essi, come, ad esempio: *Agrees With*, *Cites Cites As Authority*; *Cites As Data Source*; *Cites As Evidence*; *Cites As Metadata Document*; *Cites As Recommended Reading*; *Cites As Related*; *Cites As Source Document*; *Cites For Information*; *Compiles*; *Confirms*; *Contains Assertion From*; *Corrects*; *Credits*; *Critiques*; *Derides*; *Disagrees With*; *Discusses...*) [V. sito: *CiTO, the Citation Typing Ontology*]. Le citazioni caratterizzate dalla CITO possono essere dirette, indirette ed esplicite.

CITO offre soluzioni pratiche per la creazione delle infrastrutture citazionali (“*Citation Notification Services*”) che possono arricchire il profilo semantico dei contenuti digitali in *Repository Open Access*. Alla base del “*Citation Notification Service*” vi è il protocollo *TrackBack* che può essere applicato agli articoli accessibili in modalità *Open Access (OA)*, per accedere, elaborare e analizzare un volume crescente di dati citazionali.

L’Ontologia CITO viene citata nel documento del *Digital Curation Centre* - “*How to Cite Datasets and Link to Publications*” [V. sito], una guida che illustra la creazione di collegamenti tra pubblicazioni di ricerca e dati citazionali su cui esse si basano e/o incorporano, così da rendere possibile la navigazione tra diversi *dataset* (da un articolo ai dati citazionali e viceversa) per chi legge un articolo. Le potenzialità insite nell’esposizione di un insieme di dati citazionali ai fini della valutazione di un lavoro di ricerca, “sono oggetto di un certo interesse da parte di alcune iniziative correlate agli Archivi Aperti” [De Robbio, 2012, p.15], dal momento che le attività di ricerca si realizzano e dipendono dalla possibilità di accedere, elaborare e analizzare un volume crescente di dati correlati.

2.6.5. Extensible Observation Ontology (OBOE)

L’Ontologia “*Extensible Observation Ontology*”(OBOE) è un’ontologia formale per catturare la semantica di osservazioni e misurazioni scientifiche. Questa ontologia fornisce una base per l’aggiunta di dettagliate annotazioni semantiche ai dati scientifici, per ancorare il “significato” intrinseco di dati provenienti dalla descrizione di studi osservazionali (es. da un articolo di ricerca) [O’Brien, 2010].

La struttura di base dell’ OBOE (Figura 85) si compone di cinque classi etichettate come: “*Observation*”, “*Entity*”, “*Measurement*”, “*Characteristic*”, “*Measurement Standard*”, e sei

proprietà: “hasContext”, “ofEntity”, “hasMeasurement”, “hasValue”, “hasPrecision”, “usesStandard”, “ofCharacteristic”.

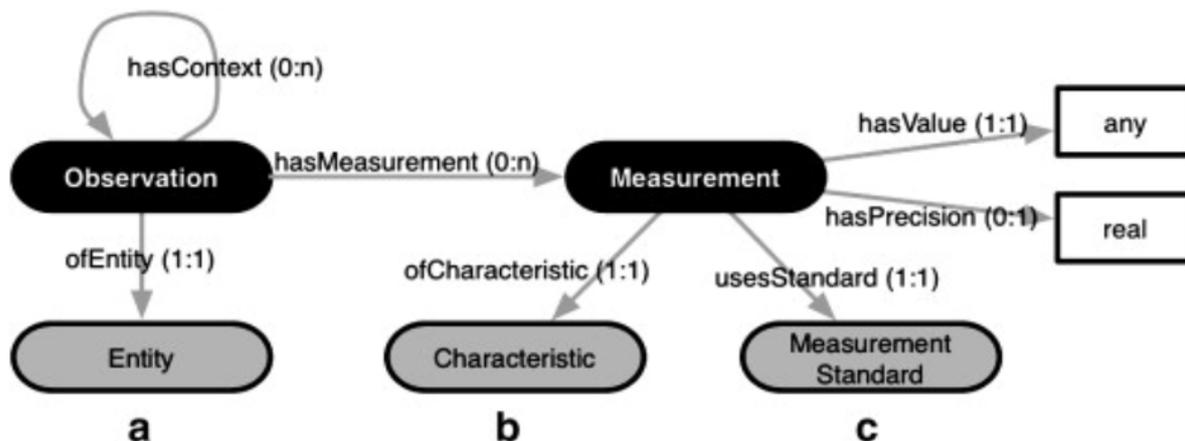


Figura 85. Modello OBOE (*Extensible Observation Ontology*): descrizione di dati e processi scientifici empirici

Proprietà aggiuntive possono essere associate alla Classe “*Measurement*”, per catturare le unità temporali e locali di una misurazione scientifica, l’unità descrivente il soggetto che ha registrato una misurazione, i protocolli di misurazione/i, e così via. Proprietà simili possono essere aggiunte anche alla Classe “*Observation*”.

2.6.6. Nano-pubblicazioni: un approccio per comunicare le entità concettuali di un prodotto di ricerca

Una nano-pubblicazione è una nuova forma di comunicazione scientifica [Castellucci, 2011] - un “*modular article*” [Castelli et al, 2012, p.33] - espressa attraverso gli elementi modulari minimi descrittivi un’informazione scientifica pubblicabile.

“*Nano-publications are a natural response to the explosion of high-quality contextual information that overwhelms the capacity of conventional research articles in scholarly communication*” [V. sito: NANOPUB.ORG].

Tali elementi modulari possono rappresentare le asserzioni su qualsiasi unità concettuale che possa essere identificata in modo univoco e attribuita al suo autore. Gli elementi modulari minimi vengono percepiti

come *piccoli moduli di informazione*, che nel loro insieme potrebbero *aggregarsi a comporre una struttura rizomatica reticolare*, per entità e secondo schemi tipologici in domini definiti [...] strutture comunicative parallele rispetto all’articolo scientifico narrativamente lineare [che] si dirigono verso un nuovo paradigma concettuale [De Robbio, 2011 b, p.24].

Gli elementi modulari di una nano-pubblicazione sono rappresentati da nodi o blocchi a supporto descrittivo di singole annotazioni (affermazioni) su una unità concettuale (es. un articolo, le citazioni, *post* o *tweet* su un articolo) (Figura 85).

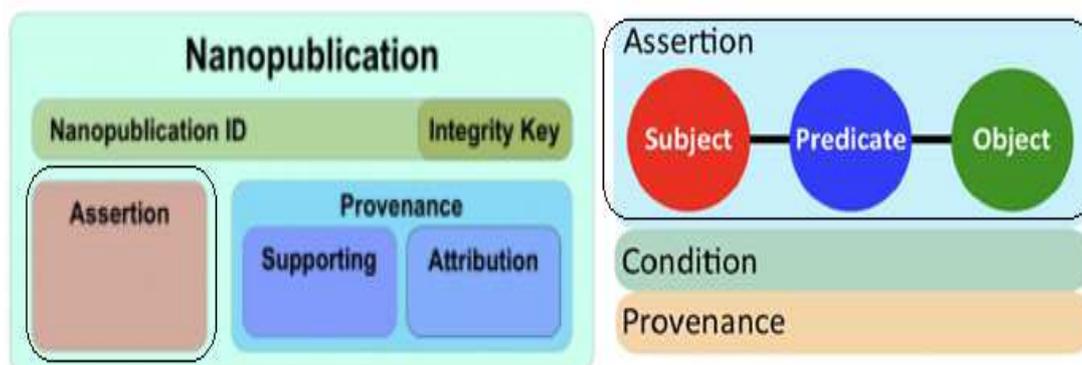


Figura 86. Le unità concettuali di una Nano-pubblicazione

Una Nano-pubblicazione ha due elementi fondamentali:

1) “Asserzione”: una unità minima di pensiero, che esprime una relazione tra due concetti “Soggetto” (es. “Retorica”) e “Oggetto” (es. “Aristotele”), usando un terzo concetto “Predicato” (es. “è scritta”).

Un’asserzione può essere serializzata utilizzando il modello RDF (*Resource Description Framework*), al fine di permettere a nano-pubblicazioni di essere “*machine readable*” e di aprire la strada ad una interoperabilità universale tra diverse asserzioni.

2) “Provenienza”: sono le informazioni descrittive (es. metadati) che forniscono un contesto per l’Asserzione”. Tali informazioni possono descrivere, ad esempio, autori, istituzioni, sovvenzioni, link agli identificatori univoci.

Ad ognuna delle affermazioni della tripla “Soggetto”, “Predicato”, “Oggetto” viene assegnato un identificatore univoco, ad esempio un indirizzo web URI (*Uniform Resource Identifier*) a sé stante, dando così il luogo al modello “nano-pubblicazione” che correla al proprio interno diverse dichiarazioni a supporto descrittivo di una certa entità (es. articolo, parte di un articolo).

Secondo questo modello, dati, immagini, suoni, simulazione e video fanno parte dell’ambiente di pubblicazione circostante il testo. Tutti questi componenti sono legati insieme e costituiscono un articolo modulare, essendo un modulo definito come una unità informativa concettuale auto-contenuta. Ciascuna unità informativa dovrebbe essere ben definita e dotata di un insieme di metadati ciascuno dei quali ne descrive un aspetto. Lo sviluppo di un articolo come aggregazione di moduli indipendenti ma interagenti fra loro è resa possibile dalla tecnologia Internet [Castelli et al, 2012, p.33].

Lo scopo di una Nano-pubblicazione è quello di fornire un meccanismo robusto per aggregare le informazioni correlate in una sola base di conoscenza (“*knowledge base*”) da cui possono essere tratte (dedotte) nuove affermazioni e create nuove inferenze di conoscenza, contribuendo così alla creazione di un reticolo di nano-pubblicazioni appartenenti a campi disciplinari diversi ma integrabili, riferite allo stesso oggetto di ricerca:

il legame tra le diverse nano-pubblicazioni non sarebbe più di tipo puramente bibliografico e nemmeno tematico, ma addirittura testuale perché i frammenti informativi sono integrabili e puntano, in qualche maniera, a completarsi [De Robbio, 2011 b, p.27].

La robustezza di una Nano-pubblicazione è dovuta alla robustezza delle proprie annotazioni (entità concettuali⁵) supportate da identificatori univoci e permanenti (es. *Life-Science Identifier* – LSID; *Digital Object Identifier* – DOI; *Uniform Resource Name* – URN) per rappresentare una risorsa nel tempo e per valutare l’attendibilità di questa o quella dichiarazione.

Con Nano-pubblicazioni è possibile disseminare i dati individuali come pubblicazioni indipendenti, con o senza articolo/i di ricerca o altri dati che le possono accompagnare. Poiché una nano-pubblicazione può essere attribuita e citata, essa fornisce incentivi per i ricercatori per

⁵ Per la categorizzazione dei concetti, ad esempio, nelle scienze naturali è stato creato un *Repository* universale di concetti ad Accesso Aperto *Concept Wiki*, <http://www.conceptwiki.org/>

rendere disponibili i loro dati in formati standard (RDF), che guidano l'accessibilità dei dati e l'interoperabilità tra i contenuti digitali.

2.6.7. *Open Archives Initiative – Object Reuse and Exchange Protocol a supporto del modello “Enhanced Publication”*

Il modello OAI-ORE (*Open Archives Initiative Object Reuse and Exchange*) [V. sito] è uno Standard per la descrizione per la descrizione e scambio di aggregazioni di risorse, o oggetti digitali composti (compound objects) sul web. Queste aggregazioni permettono di combinare risorse distribuite di vario tipo, rappresentando, ad esempio, “una pubblicazione come un “pacchetto di unità informative” che può includere testo, dati, immagini, video, audio e pacchetti software collegati fra loro mediante relazioni semantiche” (Figura 87) [Castelli et al, 2012, p.29]. L’obiettivo di queste aggregazioni è rendere disponibile il loro ricco contenuto a programmi applicativi di supporto ad attività di ricerca, scambio, visualizzazione, ri-uso e conservazione di risorse informative.

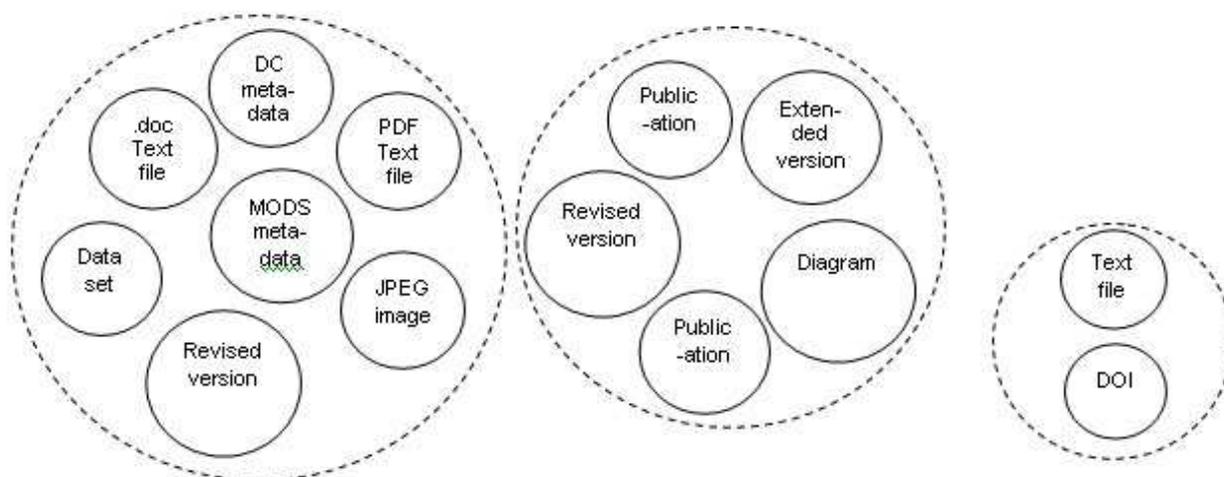


Figura 87. Possibili aggregazioni di risorse informative digitali tramite OAI-ORE (*Open Archives Initiative Object Reuse and Exchange*)

Il modello OAI-ORE (Figura 88) si esprime tramite i meccanismi:

- 1) “*resource map*” (mappa delle risorse) per indicare e descrivere la provenienza di metadati che descrivono oggetti composti sul web;
- 2) “*proxies*” per descrivere ogni oggetto composto, dalla prospettiva di una specifica aggregazione a cui tale oggetto può appartenere/relazionarsi.

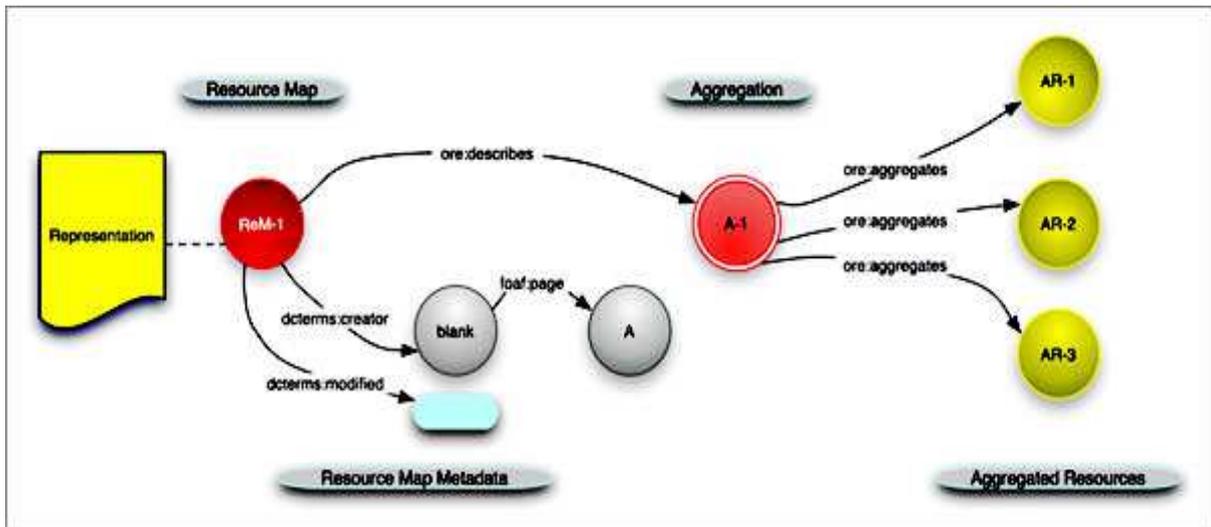


Figura 88. Modello OAI-ORE: dalle “Risorse aggregate” alla “Mappa delle Risorse” [Hogenaar, 2009]

L’OAI-ORE è basato sull’architettura del World Wide Web ed è costruito sulle triple del modello RDF impiegato nei recenti sviluppi di alcune aree del Web Semantico come *Linked Data* e *Cool URIs*.

L’impiego dell’OAI-ORE negli Archivi di contenuti digitali di ricerca può avere l’obiettivo di supportare con efficacia l’interoperabilità e l’integrazione fra dati e pubblicazioni. Questo approccio può anche rendere possibile lo sviluppo di nuovi punti di accesso ai documenti, oltre a quelli creati tramite la catalogazione descrittiva tramite i metadati, nonché la trasformazione di Archivi digitali in sistemi altamente interoperabili e applicazioni di servizi a valore aggiunto [Tarrant, 2009].

Gli effetti del protocollo OAI-ORE possono essere osservati nei Servizi di *Citation Tracking* in JSTOR [V. sito], nonché nel Portale di ricerca dei contenuti scientifici NARCIS [V. sito] (Figura 89) dove OAI-ORE è impiegato per la formazione di cosiddetti “*Enhanced Publications*” o pubblicazioni potenziati.



Figura 89 “*Enhanced Publications*” accessibili via il Portale e l’aggregatore dei contenuti di ricerca NARCIS

Il modello “*Enhanced Publication*” rappresenta

una pubblicazione dotata anche di dati a supporto dei risultati da essa riportati. Questi dati possono anche includere informazioni di tipo contestuale, oppure di provenienza dei dati. Tutti questi componenti sono esplicitamente collegati all’interno di un singolo pacchetto che costituisce, quindi, un oggetto digitale composito, contenente una varietà di componenti eterogenei collegati fra di loro, eventualmente distribuiti in diverse ubicazioni [Castelli et al, 2012, p.33].

La “*Enhanced Publication*” e “*modular article*” sono due esempi significativi di pubblicazioni di oggetti compositi (*compound objects*).

Entro il progetto OpenAIREplus (*Open Access Infrastructure for Research in Europe*) [V. sito], finanziato dalla Commissione Europea, sono stati sviluppati alcuni servizi dimostrativi di “*Enhanced Publications*” per i domini di ricerca:

- “*Life Sciences*”,
- “*Social Sciences*”,
- “*Humanities (Oral History)*”

[V. sito: *OpenAIRE's 'Enhanced Publications' Pilots*] (Figura 90) che collegano diverse entità informative a supporto delle pubblicazioni di ricerca attraverso le relazioni tra pubblicazioni, progetti, e dati di ricerca.

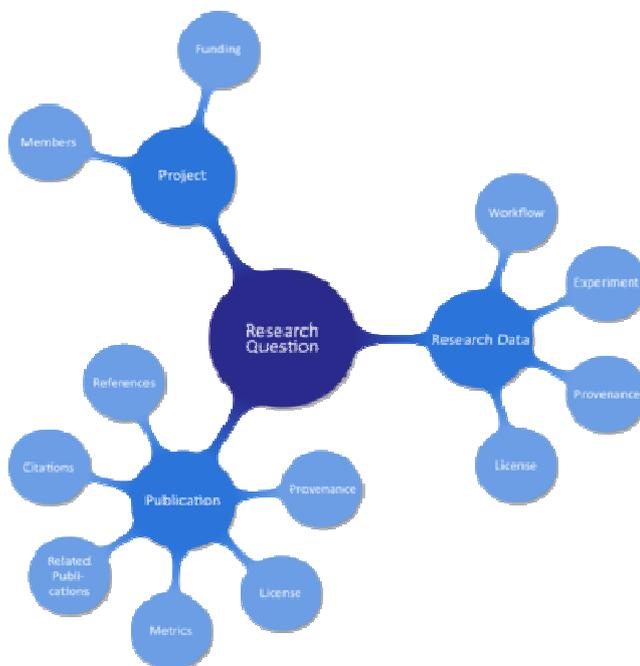


Figura 90. *OpenAIRE's 'Enhanced Publications' Pilots*

2.6.8. Approcci di aggregazione e interconnessione di oggetti digitali complessi nella Biblioteca digitale europea *Europeana*

Europeana [V. sito] è un progetto (con tre sotto-progetti: “*Europeana Connecting Project*”, “*Europe's cultural collections*”, “*Professional Knowledge-Sharing Platform*”) della biblioteca digitale europea che riunisce contributi digitalizzati di profilo culturale dalle istituzioni dei 27 Paesi dell’Unione Europea. Il progetto Europeana è ormai completamente integrato nell’Agenda Digitale Europea.

Il progetto Europeana è stato promosso dalla Commissione Europea che ha proposto di digitalizzare il patrimonio culturale di tutte le istituzioni culturali dell’Unione e di metterlo a disposizione del pubblico attraverso il Portale Europeana – un servizio web di alta qualità - che attualmente contiene più di 23 milioni di contenuti, tra immagini, video, file audio e testi.

Dal 1° luglio 2012 Europeana presenta i metadati (informazioni descrittivi) aggregati da diversi Archivi, Musei e Biblioteche esclusivamente sotto forma di *Linked Open Data* (LOD), cioè come dati aperti e connessi, contribuendo così all’interoperabilità tra i contenuti digitali posseduti.

La pubblicazione di metadati come LOD è resa possibile grazie alla sottoscrizione da parte dei partner - che includono sia singoli *Content Provider* che aggregatori di *Content Provider* - di Europeana dell’accordo sullo scambio dei dati (*Data Exchange Agreement:DEA*) [V. sito]. Secondo DEA, i partner devono fornire a Europeana almeno un minimo insieme di informazioni

(metadati) a ogni materiale che intendono condividere con Europeana e rilasciare il permesso di pubblicare tali metadati sotto la Licenza Aperta Creative Commons 0 (CC0 1.0 Universal Public Domain Dedication). Sostenendo la forza della CC0, Europeana include l'autorizzazione per il riutilizzo dei contenuti aggiunti da Europeana da parte di terzi. In questo modo Europeana può rendere i contenuti culturali non solo fruibili, ma anche facilmente scaricabili, modificabili, riutilizzabili dagli utenti della rete.

Attualmente Europeana ha a propria disposizione 4 milioni di registri di metadati (*metadata records*) forniti dai partner e rilasciati sotto la CC0.

Anche gli identificatori come URL (*Uniform Resource Locator*) - che sono a supporto identificativo di locazione di contenuti aggregati da Europeana - sono rilasciati/pubblicati sotto la licenza CC0. Attraverso questi URL i contenuti (es. immagini) possono essere liberamente visualizzati tramite il Portale Europeana.

As soon as an image is published at a publicly accessible URL (on Europeana or Content Provider webpages), the URL is published and can be used by anyone to display that image, by simply entering the URL into a browser or similar application [...] Since the preview will be displayed on europeana.eu, a clever person can still link to it, but it will be much harder to build applications that automatically provide access to a large number of previews. In any case, the preview keeps the rights assigned by the Content Provider and any unauthorized use of it may be persecuted by law

[V. sito: *Use of Content in Linked Heritage and Europeana*, p. 6].

Dietro *Europeana* si nasconde la sfida più grande per il patrimonio culturale europeo: creare un sistema di metadati unico in tutta l'Unione. Intanto, come uniformare tra loro i contenuti di istituzioni diverse per renderli accessibili a tutti e rintracciabili con una semplice ricerca web? Attraverso un unico modello specifico di ricerca semantica, l'EDM (*Europeana Data Model*) [V. sito].

L'EDM è costituito attraverso il riutilizzo dei modelli di interoperabilità di metadati come RDF, RDFS, OAI-ORE, SKOS e DCMI Terms. Dal punto di vista pratico, l'attuazione del modello EDM è possibile attraverso le dettagliate Linea Guida:

- 1) “*Europeana Data Model Primer*”;
- 2) “*Europeana Data Model Mapping Guidelines*”.

Nel documento “*Europeana Data Model Fact Sheet*”, diretto ai cosiddetti *Policy makers* di partner di Europeana, si descrive come i loro dati possano essere arricchiti attraverso il riutilizzo di altri fonti terze, segnalando una base di conoscenza aperta sul web DBpedia che per qualche ragione appare al centro del famoso diagramma in forma di nuvola di LOD. In quel documento risultano inoltre chiaro come l'EDM contribuisca alla rete di dati o Web Semantico, e come si debba implementare l'EDM, una struttura equilibrata che permette di ospitare diversi modelli di metadati esistenti e quelli emergenti del mondo di Biblioteche, Archivi, e Musei.

2.6.9. Raccomandazioni “LODE-BD” (*Linked Open Data – enabled Bibliographic Data*)

Le Raccomandazioni “LODE-BD” (2.0) [V. sito] promosse dal 2012 dalla comunità web AIMS “*Agricultural Information Management Standards*” riguardano le raccomandazioni per lo sviluppo della realtà pratica - *Linked Open Data-enabled bibliographic data*, ossia dati bibliografici (collezioni di dati o metadati descrittivi una risorsa) abilitati come *Linked Open Data* (LOD).

Le Raccomandazioni si concentrano sugli indirizzi:

- *metodologico*, ovvero organizzazione concettuale degli elementi di metadati per una migliore rappresentazione di risorse informative e
- *tecnico*, ovvero le strategie per la codifica dei valori di metadati utili alla creazione delle entità LOD, in corrispondenza con le esigenze di affiancamento di elementi di metadati con gli strumenti di *Authority Control*, di arricchimento semantico, con gli appropriati

Standard e le norme per la compilazione dei metadati, al fine di contribuire ad una maggiore comunicazione tra contenuti e, di conseguenza, per creare le basi solide per l'interoperabilità tra le infrastrutture che li ospitano.

“LODE-BD” sottolineano, intanto, che per provvedere ad una migliore interoperabilità tra i contenuti reale ed effettiva occorre renderli Aperti (*Open*), quindi è necessario "aprirli" tramite l'applicazione di Licenze e formati aperti.

Le questioni correlate alle strategie di codifica di metadati a vario livello sono:

1) Quali *Standard* (riferiti a modelli, profili) di *metadati* e *Vocabolari* devono essere seguiti al fine di accentuare la correttezza, l'autenticità, la semantica, l'interoperabilità nel Web Semantico dei valori di metadati?

La risposta di “LODE-BD” è quella di:

- seguire le indicazioni di Standard, Schemi e Vocabolari ampiamente riconosciuti ed adoperati nella comunità di riferimento;
- identificare (ove e' possibile) i contenuti e le loro parti tramite gli identificatori URIs (*Uniform Resource Identifiers*);
- adoperare Schemi e Vocabolari disponibili in modalità *Linked Open Data* nonché quelli Schemi o Ontologie (es. *Schema.org*, *BIBO Ontology*) capaci di integrare meglio e di esporre i contenuti (oggetto della descrizione) nel Web Semantico;

2) Qual è l'insieme minimo di metadati su cui i contenuti digitali devono appoggiarsi per essere significativi entro un contesto di condivisione?

A proposito “LODE-BD” propongono un insieme di metadati distribuiti in nove Categorie:

1. *TITLE INFORMATION* (*Title/Alternative title*);
2. *RESPONSIBLE BODY* (*Creator, Contributor, Publisher*);
3. *PHYSICAL CHARACTERISTICS* (*Date, Identifier, Language, Format/Medium, Edition/Version, Source*);
4. *HOLDING/LOCATION INFORMATION* (*Location/Availability*);
5. *SUBJECT INFORMATION* (*Subject*);
6. *DESCRIPTION OF CONTENT* (*Description/Abstract/Table of Contents; Type/Form/Genre*);
7. *INTELLECTUAL PROPERTY RIGHTS* (*Right Statements*);
8. *USAGE* (*Audience/literary indication/education Level*);
9. *RELATION*

3) Qual è la metodologia migliore per definire il profilo concettuale di ogni metadato?

In relazione “LODE-BD” offrono l'uso di “ALBERI DECISIONALI” (Figura 91).

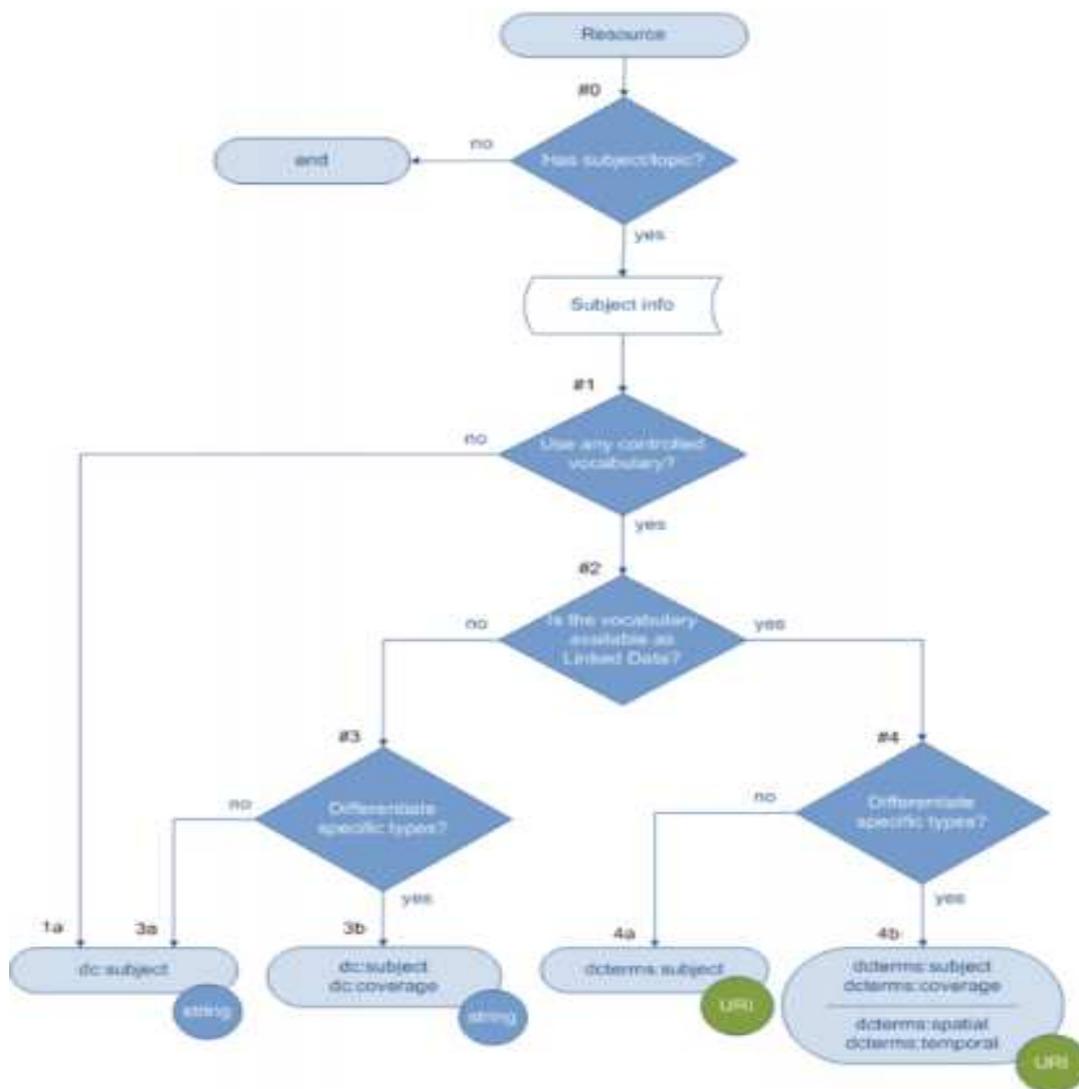


Figura 91. Albero decisionale “LODE-BD” per la scelta delle decisioni sulla concettualizzazione del profilo del metadato SOGGETTO

Gli Alberi decisionali rappresentano *diagrammi di flusso* per fornire l’assistenza ai gestori delle infrastrutture di contenuti digitali nel processo di *individuazione delle proprietà* rilevanti di *metadati* e per facilitarne la selezione delle *strategie* di codifica appropriate nelle più svariate situazioni ed esigenze locali.

Le Raccomandazioni “LODE-BD” sono uno strumento utile per i professionisti dell’informazione, per i gestori dei *Repository* e Biblioteche digitali, dal momento che forniscono i punti d’azione più coerenti alla *scelta e codifica* dei *metadati* in possesso, cosicché i produttori dei dati possano costruire *sistemi coerenti* alla sincronizzazione (armonizzazione), condivisione e alla collaborazione dei medesimi dati, abilitando nel contempo la comunicazione di questi ultimi nel contesto del Web Semantico.

Non è più solo importante fornire i dati in un *formato aperto*, ma risulta fondamentale munire i dati degli *step utili* a comprendere gli *sviluppi* che sono intervenuti su tali dati, rendendo anche graficamente i processi intervenuti sui dati stessi tramite visualizzazioni in *mappe concettuali* e *diagrammi di flusso*. Se tutto il processo risulta trasparente, il *dato* può essere *riutilizzato* nelle sue singole fasi, tramite *link adeguati* [De Robbio, 2011 a, p.19].

2.7. Progetto dell'Archivio di ricerca *Open Access* VOA3R

Il VOA3R (*Virtual Open Access Agriculture & Aquaculture*) *Open Access* Repository [V. sito] è un Progetto di ricerca Europeo finanziato (2010 – 2013) dal Settimo Programma quadro dell'Unione Europea (*ICT Policy Support Programme - Theme 4 – Open Access to Scientific information*).

L'obiettivo principale di VOA3R è quello di migliorare la disseminazione dei risultati della ricerca europea nei campi scientifici di Agricoltura e Acquacoltura, utilizzando approcci innovativi per la comunicazione e la condivisione dei prodotti della ricerca. Il risultato del Progetto è lo sviluppo di servizi ad integrazione degli Archivi (un approccio di ricerca federata dei contenuti di più Archivi da un solo Archivio digitale) e dei sistemi di pubblicazione, attraverso applicazione delle Raccomandazioni "LODE-BD" (*Linked Open Data – enabled Bibliographic Data*).

VOA3R ha creato un Portale che offre un unico punto d'accesso per la ricerca, l'accesso, la consultazione e la diffusione dei materiali di ricerca che provengono da una rete federata di 14 *Repository Open Access*. Attualmente, da VOA3R è possibile accedere a più di 500.000 pubblicazioni ad Accesso Aperto.

In quanto all'infrastruttura tecnica del Portale VOA3R, essa si poggia su una rete globale di dominio pubblico denominata AGRIS (*International System for Agricultural Science and Technology*) nella quale, allo stato attuale, sono presenti più di 4 milioni record bibliografici (o schede di metadati) relativi a contenuti di ricerca prodotti nei campi scientifici di Agricoltura e Acquacoltura. VOA3R ha esteso le funzionalità della rete AGRIS per mezzo di componenti aggiuntivi, tra cui i nuovi servizi di ricerca potenziati dall'impiego di:

- Thesaurus online pubblicato in modalità *Linked Open Data* - AGROVOC cui valori sono identificati dagli indirizzi URI appoggiati sul protocollo Web Http,
- nuovi modelli e funzionalità del *Social Network* (Web 2.0) di comunicazione attiva in rete [Le Hénaff, 2012].

2.7.1. Programmazione dei requisiti di metadati estensibili di qualità

Durante il processo di sviluppo del Profilo applicativo di metadati per l'Archivio digitale VOA3R, il Gruppo di lavoro ha individuato e documentato diversi elementi di metadati e loro requisiti in conformità alle funzionalità avanzate per la descrizione e ricerca dei contenuti posseduti dall'Archivio.

Il profilo di metadati VOA3R – che può fungere da modello per i profili di metadati al servizio degli Archivi Istituzionali *Open Access* – si prefigge i seguenti obiettivi:

- 1) facilitare la descrizione, la condivisione e il riutilizzo dei contenuti digitali di ricerca (qui attinenti al campo dell'Agricoltura) in Archivi ad Accesso Aperto;
- 2) sostenere la scoperta delle risorse della ricerca;
- 3) sostenere l'identificazione e la classificazione delle risorse utilizzando esistenti Ontologie e Tassonomie;
- 4) abilitare l'interconnessione semantica delle risorse sostenendo l'esposizione di metadati in formati di *Linked Open Data*;
- 5) combinare metadati bibliografici con i dati di ricerca del dominio della comunità di ricerca;
- 6) servire come un formato di interscambio tra metadati di vari Archivi (*Repository*) usando diversi Standard e formati di metadati;
- 7) raccogliere (catturare) i metadati degli Archivi dopo il controllo di qualità dei loro valori;
- 8) supportare i servizi a valore aggiunto per la comunità di ricerca.

A seguire, alcuni requisiti associati ai principi di progettazione metadati di alto livello, riflettendo principalmente la loro natura e le caratteristiche del Profilo di metadati risultanti (Tabella 21).

1.	<p>Ampio riuso di modelli/profilo di metadati esistenti (standard) Nella pratica corrente, la raccolta o cross-ricerca dei contenuti da più <i>Repository</i> deve far fronte a un numero di questioni legate ai metadati. Tra questi vi è una questione importante che riguarda l'uso dello schema di metadati descrittivi Dublin Core/DC semplice (basato su 15 elementi di base) – il cui uso è considerato fondamentale per la raccolta da parte di Archivi aggregatori (<i>Service Provider</i>) – attraverso il protocollo OAI-PMH - di metadati da Archivi Istituzionali/disciplinari <i>Open Access</i>. Lo sviluppo e l'implementazione da parte degli Archivi di ricerca di un Profilo di metadati ricche basato sugli schemi e sui profili di metadati esistenti ampiamente riconosciuti compatibili con il DC semplice (es. per gli Archivi Istituzionali, il profilo di metadati SWAP/<i>Scholarly Work Application Profile</i>) - consentirebbe a <i>Service Provider</i> di offrire servizi di ricerca ricche e a diverso valore aggiunto.</p>
2.	<p>Metadati come <i>Linked Data</i> Sempre più set di dati delle risorse, Vocabolari controllati e <i>Authority File</i> sono disponibili come <i>Linked Data</i>, ed esprimono le loro entità concettuali con identificatori web (Http://) URIs. Il loro uso è importante nell'abilitare i metadati in Archivi digitali a diventare <i>Linked Data</i>, con il successivo potenziale di essere ampiamente collegati con altri dati di qualità sul web.</p>
3.	<p>Metadati consistenti È necessario che ogni Archivio digitale non solo sviluppi e implementi il Profilo di metadati consistente (V. il punto 1 della presente Tab.), ma anche che condivida le informazioni sul medesimo Profilo con gli Archivi aggregatori (<i>Service Provider</i>), cosicché questi ultimi possano sviluppare le adeguate mappature tra diversi Profili di metadati raccolti da <i>Data Provider</i>, evitando così l'appiattimento dei metadati raccolti ai metadati DC semplici (non qualificati).</p>
4.	<p>Metadati estensibili Un profilo di metadati sviluppato e condiviso tra diversi Archivi può essere arricchito con elementi di maggiore espressività, in termini di entità, attributi e relazioni che ricadono sotto la competenza più ampia di processi di ricerca e di contesto. Sarebbe opportuno di sviluppare i metadati estensibili (es. la mappatura tra diversi Profili di metadati) attraverso il reciproco sostegno da parte di responsabili di vari Archivi cooperanti</p>
5.	<p>Coinvolgimento di metadati nei servizi a valore aggiunto dell'Archivio Sarebbe opportuno utilizzare i metadati programmati dell'Archivio a supporto di servizi avanzati come:</p> <ul style="list-style-type: none"> 5.1) la rappresentazione del ciclo di vita (<i>lifecycles</i>) di contenuti di ricerca. Occorrono i metadati descrittivi lo stato (es. data di revisione, data di pubblicazione; Versione iniziale, intermedia, finale; Versioni tradotti ecc.) di una risorsa, sia all'interno (che all'esterno) di un Archivio. Ciò è possibile tramite l'uso di metadati che esprimono le relazioni (es. "<i>HasVersion</i>", "<i>HasTranslation</i>", "<i>ConformsTo</i>"). 5.2) la valutazione di ricerca scientifica. è opportuno provvedere alla mappatura tra il profilo di metadati di un Archivio con lo Standard a supporto di sistemi di ricerca CERIF (V. § 2.6.2. del presente Cap.); 5.3) la ricerca semantica di contenuti, basata sull'uso di vocabolari controllati semantici (Thesauri, Classificazioni) entro gli Archivi; 5.4) Ricerche incrociate tra i contenuti dell'Archivio e quelli del Catalogo della

	Biblioteca dell'istituzione. Il profilo di metadati adottato per un Archivio Istituzionale dovrebbe essere compatibile e comunicabile con i principi e gli approcci della catalogazione della Biblioteca.
6.	<p>Qualificazione delle proprietà di metadati tramite l'uso di valori normalizzati dai documenti di autorità come "Name Authority" e Vocabolari Controllati semantici</p> <p>È importante:</p> <p>6.1) Provvedere all'uso delle forme autorevoli per i Nomi: qualificare le proprietà di metadati descrittivi i nomi di Autori, Istituzioni, Luoghi geografici;</p> <p>6.2) Classificare le tematiche delle risorse tramite i valori di vocabolari semantici controllati come Thesauri, Classificazioni, Soggettari ampiamente riconosciuti e implementati nei diversi Archivi.</p>
7.	<p>Metadati per rappresentare le "risorse complesse"</p> <p>Una "risorsa complessa" può essere creata attraverso una sintesi di altre risorse, o tramite una raccolta o sequenza di altre risorse correlate nel contesto, ad esempio:</p> <ul style="list-style-type: none"> • un libro che include una collezione o sequenza di pubblicazioni; • una Tesi di dottorato che include una serie di grafici e dati di ricerca (<i>research data</i>); • un manoscritto basato su una collezione di fotografie; • un articolo che usa determinati metodi o strumenti di ricerca; • un rapporto di ricerca collegato ad un progetto; • un articolo che critica i risultati esposti in un altro articolo. <p>Al campo concettuale di una "risorsa complessa" appartiene anche una "ENHANCED PUBLICATION" (V. § 2.6.7. del presente Cap.), un "LEARNING OBJECT".</p> <p>Un "Learning Object" è una composizione di diversi oggetti digitali e ha le seguenti proprietà:</p> <ul style="list-style-type: none"> • contiene le relazioni/i collegamenti tra oggetti/risorse di appartenenza; • contiene una tabella di oggetti di appartenenza (contenuti); • ha un appropriato set di metadati; • ha un appropriato identificatore permanente (PID), • ha un appropriato formato definito per la visualizzazione. <p>I collegamenti entro le "risorse complesse" devono essere supportati attraverso i metadati che esprimono diverse annotazioni (es. "CritiquesOn", "Critiques", "Supports", "CommentsOn", "CommentsOn", "isCommentedBy", "Extends").</p>
8.	<p>Metadati a supporto di contenuti di ricerca provenienti dal Web 2.0</p> <p>Elementi e valori di metadati devono essere in grado descrivere le risorse di ricerca correlate che provengono da fonti Web2.0 (es. <i>YouTube</i>, <i>SlideShare</i>, i blog di ricerca, <i>feed</i> di <i>Twitter</i>, ecc.).</p> <p>La qualità di tali risorse deve essere precedentemente valutata dagli esperti umani. La loro successiva indicizzazione e metadattazione deve essere inclusa nella collezione dell'Archivio (es. in forma di un <i>stream</i>) e correlata con i metadati di risorse depositati che richiamano le stesse o simili attività di ricerca,.</p>
9.	<p>Metadati che collegano la rappresentazione di uno stesso contenuto di ricerca in diversi Archivi (Istituzionale, Disciplinare, Personale, dell'Editore)</p> <p>Una risorsa di ricerca può essere creata (pubblicata, presentata) da vari tipi di fornitori di contenuto (<i>Content Providers</i>). Sarebbe opportuno esprimere - attraverso metadati - i collegamenti tra questi fornitori in relazione alla creazione della medesima risorsa.</p>
10.	<p>Metadati a supporto di navigazione tra diverse risorse</p> <p>I metadati che descrivono le relazioni (V. il punto 7 della presente Tab.) devono essere incorporati nell'interfaccia di ricerca , al fine di essere visibili e permettere la</p>

	navigazione tra le risorse in ogni direzione programmata.
11.	Attivare la distinzione tra metadati generati automaticamente dalla macchina/sistema e quelli creati/programmati manualmente Entrambe le tipologie di metadati necessitano la validazione della loro qualità da parte di professionisti dell'informazione.
12.	Ad ogni profilo di metadati adottato in un Archivio deve essere attribuito il proprio identificatore univoco permanente (PID).

Tabella 21. Requisiti per la programmazione dei requisiti di metadati estensibili di qualità (il caso dell'Archivio VOA3R)

2.7.2. Costruzione del sociale (Web 2.0) e semantico (*Linked Data*) per la comunicazione dei contenuti digitali

In questo paragrafo è fornita una veloce presentazione delle funzionalità avanzate sviluppate entro il Progetto VOA3R. A tal fine è fornita una panoramica degli strumenti utilizzati nella creazione del *Repository Open Access* VOA3R per completare la granularità dell'uso del materiale di ricerca attraverso diverse funzionalità di navigazione e di valutazione per mezzo:

- dell'approccio consistente, formulato mediante l'unione dei meccanismi del *Social* (Web 3.0) e *Semantic Web* (Web 3.0, *Linked Data*) e
- di metriche valutative di metadati incorporati nel *Repository* per contribuire col valore aggiunto alla diffusione strutturata e partecipativa del materiale di ricerca in rete.

Attraverso *Social networking* (Web 2.0), nel Portale VOA3R è possibile gestire diverse comunicazioni virtuali attive (attraverso il *microblogging*, *social bookmarking*) tra i gruppi di utenti registrati, formati da accademici, ricercatori, studenti, e professionisti del campo, offrendo anche la possibilità di commentare e valutare i contenuti di ricerca attraverso le annotazioni, revisioni, e le funzionalità di *rating* e 'Like' per ogni *item* (incluso il materiale audio-visivo) depositato. L'uso di funzionalità *Feed* (aggiornamenti sui nuovi contenuti) del Web 2.0 può essere di un certo interesse non solo in termini di trasparenza di contenuti di ricerca raccolti nell'Archivio, ma anche nella manipolazione stessa di questi contenuti, inteso come bene comune.

Ogni utente registrato può creare entro il *Repository* VOA3R il suo profilo personale in cui, a parte le funzionalità appena descritte, vi sono anche dei moduli attraverso i quali è possibile dichiarare il proprio coinvolgimento in progetti, attività di ricerca istituzionali e reti sociali virtuali creati entro la *Community* VOA3R.

Il *Repository* VOA3R "also allows users to enter *Personal Contact and History Data*, which is managed alongside their research contributions, so their peers and other users can contact them online or offline for discussions and to share ideas" [Stracke et al., 2012, p.17]. Ogni utente registrato può creare una nuova comunità (gruppo) di utenti per la condivisione di interessi di ricerca simili dove poter riunire parti interessate da diversi luoghi geografici, accorciando così anche le distanze comunicative fra i membri della *Community*. Nel frattempo, tutti gli utenti del VOA3R sono diventati contributori attivi al "dialogo tra pari" nella rete professionale virtuale AIMS (*Agricultural Information Management Standards*) [V. sito]. Una simile creazione di canali comunicativi entro un unico spazio rappresenta, di certo, il valore aggiunto per il potenziamento della diffusione dei contenuti del *Repository* VOA3R il quale, tra l'altro, viene integrato nella cornice a supporto delle informazioni sulle attività promotori del movimento *Open Access* (Figura 92).

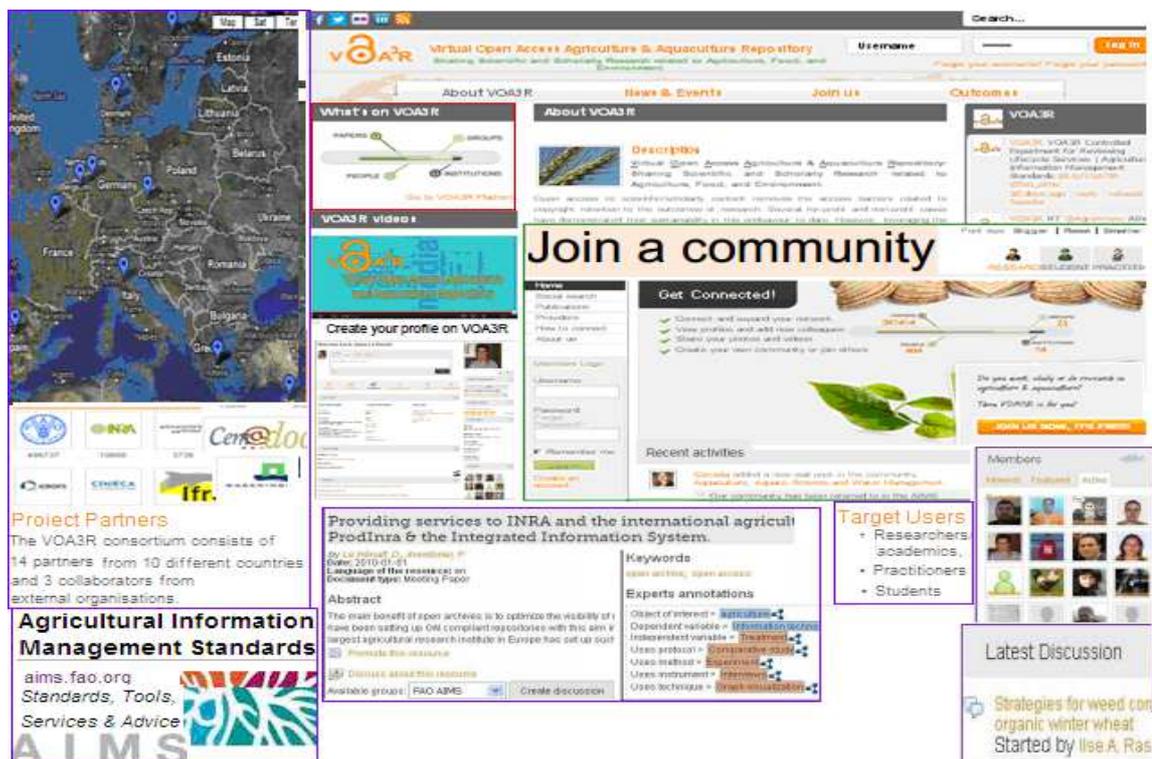


Figura 92. Portale VOA3R: partner, contenuti, servizi e modelli sociali a supporto della diffusione dei contenuti di ricerca

Già allo stato attuale di sviluppo (in modalità *beta-version*) il Portale VOA3R può rientrare tra le *Best Practices* per la creazione delle funzionalità avanzate in *Open Access Repositories*, in quanto non solo è implementato secondo l'approccio standardizzato di *Current Research Information Systems* CRIS [Keith, Asserson, 2008] (basandosi sul modello concettuale CERIF 1.3), ma anche mette in sinergia diversi contenuti e funzionalità dell'infrastruttura di ricerca attraverso i meccanismi del Web 2.0.

Inoltre, la diffusione dei contenuti di ricerca del *Repository* non si basa su un solo ancoraggio al livello dell'interoperabilità sintattica tramite l'*Open Archive Initiative Protocol for Metadata harvesting* (OAI-PMH), bensì viene anche intercorrelata per via dell'interoperabilità semantica, sfruttando le potenzialità offerte dall'ontologia basata su più di 40.000 concetti presentati in 22 lingue del Thesaurus AGROVOC, dall'*EuroCRIS Ontology server*, nonché dalle tecnologie *Linked Data* (Web 3.0) (Figura 93).

In the next year, the VOA3R partners want to push the platform development more in the direction of Linked Open Data including advanced interlinking features and multi-vocabulary support for navigational search. Finally, VOA3R establishes a European Open Access Association: For a future of research with unlimited openness [Stracke, 2012, p.17]

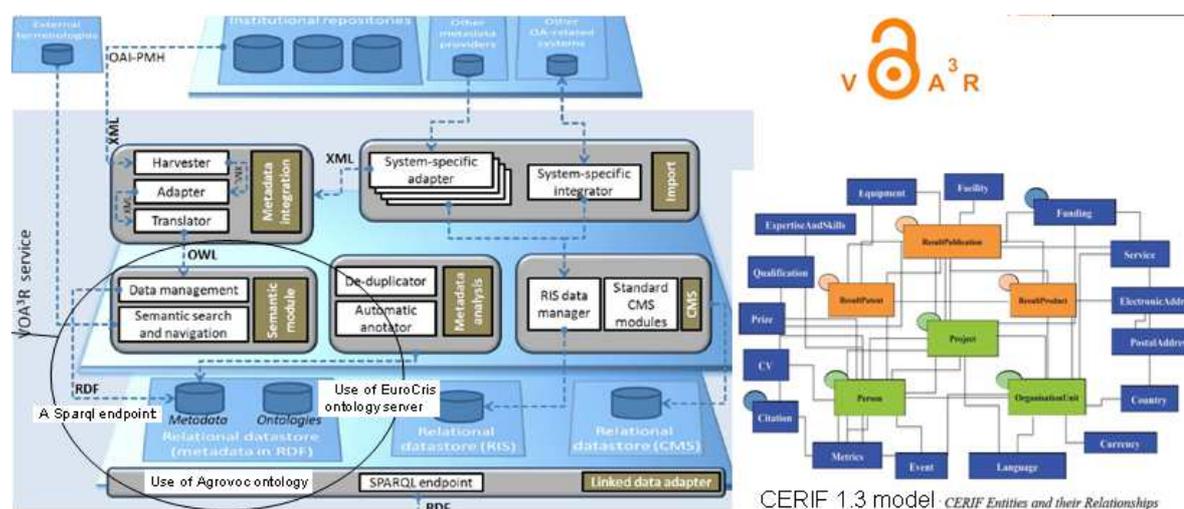


Figura 93. Infrastruttura tecnologica del *Repository* di ricerca *Open Access* - VOA3R in vista del potenziamento dei canali comunicativi dei contenuti

L'infrastruttura tecnologica del VOA3R potenziata dalle funzionalità avanzate del *Social e Semantic Web*, rinvigorisce in modo significativo l'approccio sociale aperto verso la diffusione dei contenuti di ricerca *Open Access* in rete.

2.7.2.1. Funzionalità avanzate di ricerca e di valutazione della qualità di contenuti depositati

Il *Repository* VOA3R offre una potente infrastruttura per la ricerca e navigazione dei contenuti aggregati da diversi *Repository* partecipanti, in quanto usa un approccio terminologico integrato che permette di indicizzare/taggare tutte le entità depositate attraverso la terminologia standard offerta dal sistema di organizzazione della conoscenza Thesaurus AGROVOC. Le voci controllate (voci d'autorità) di quest'ultimo sono strutturate per descrivere non solo le diverse categorie di contenuti scientifici, ma anche i nomi di persone e istituzioni. Di recente l'AGROVOC è stato pubblicato in modalità *Linked Data* con l'impiego dei formalismi di *Simple Knowledge Organization System (SKOS)* e reso visibile graficamente nella *Linked Open Data Cloud*, un insieme di cataloghi di *dataset* pubblicati sul Web come *Linked Data*, che contengono collegamenti di dati ad altri insiemi di dati collegati:

Today AGROVOC is a SKOS-XL concept scheme published as Linked Open Data cloud, containing links (as well as backlinks) and references to many other Linked Datasets in the LOD cloud [...] Searches for content in VOA3R rely on unconstrained availability, so the matter of Open use is fundamental
 [Caracciolo et al., 2012, online, non impaginato].

Oltre alle opzioni standard offerte nel VOA3R per la navigazione dei contenuti, attraverso i link a supporto delle categorie "Data provider", "Lingua", "Collezione", "Autore", il *Repository* offre le funzionalità di ricerca e navigazione avanzate come:

- "Tag Cloud" (ricerca dei contenuti in base alle parole chiavi di maggiore occorrenza, controllate attraverso il Thesaurus AGROVOC),
- "Drag and Drop" (ricerca in base alle parole chiavi applicati ai domini concettuali taggati dall'AGROVOC),
- "Timeline views",
- "Map views" (Figura 94).



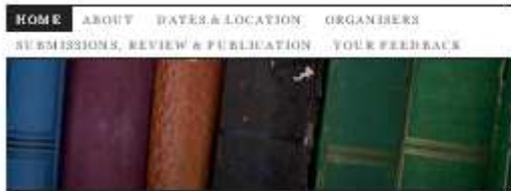
Figura 94. Funzionalità di ricerca e navigazione dei contenuti nel *Repository* VOA3R

La terminologia del Thesaurus AGROVOC è stata impiegata nel *Repository* sia allo scopo di diversificare e espandere le funzionalità di navigazione dei contenuti attraverso l'approccio "berrypicking"⁶, con il successivo filtro dei contenuti impostato per le aree spazio-temporali, che per poter esprimere meglio le intenzioni di ricerca dell'utente attraverso l'approccio di *micro-blogging*, il quale offre la possibilità di attivare comunicazioni virtuali intorno a diversi domini concettuali di ricerca appoggiati su aree semantiche formalizzate dal medesimo Thesaurus. Tale approccio consistente del VOA3R per l'organizzazione e la ricerca dei contenuti, incluso il potenziamento di rappresentazione di questi ultimi offerto dalle strategie LODE-BD (*Linked Open Data enabled Bibliographical Data*) [V. § 2.6.9. del presente Cap.] può essere considerato come un'estensione di qualità delle distribuzioni *Open Archive Initiative* che è possibile adottare in qualsiasi altra piattaforma di Archivi Aperti.

Per poter arricchire ulteriormente le funzionalità del *Repository* a supporto della qualificazione dei contenuti di ricerca ospitati, entro il progetto VOA3R è stato lanciato l'esperimento "*Controlled Experiment for Reviewing Lifecycle Services*" [V. sito] con la creazione di un apposito Blog (Figura 95).

⁶ "In 1989 Bates introduced BERRY-PICKING as a powerful metaphor highlighting that searching is not a linear process but is more accurately described as an activity that is distributed in time and space. From an information seeking perspective there are interesting similarities between berrypicking and real world wayfinding", Lueg, Christopher, Bidwell, Nicola. *Berrypicking in the Real World: A Wayfinding Perspective on Information Behavior Research*, *Proceedings of the American Society for Information Science and Technology, Volume 42, Issue 1, 2005, p.1., <http://onlinelibrary.wiley.com/doi/10.1002/meet.14504201241/full>*

VOA3R Controlled Experiment for Reviewing Lifecycle Services



NOVEMBER 10, 2012, 11:05 AM

Data-driven Schema Matching in Agricultural Learning Object Repositories

As the wealth of structured repositories of educational content for agricultural object is increasing, the problem of heterogeneity between them is becoming a serious level in learning more persistent. Ontology matching is a technique that helps to identify the correspondences in the description schemas of

You can view/ download the submitted paper from [HERE](#).

After reviewing the paper you can [leave your comment](#) on this post, and/or rate the paper (Rate this - above the title of the post).

Share this: [Leave a Comment](#)

Like this: [Like](#)

Recent Comments

[View of comments by V.V.V.V.V.](#)
[Professione Civile on this](#)

[Metadatos on this](#)

Twitter Updates

International Customization Computing of Systems, Barcelona, Spain, 7-12/2012. #Metadatos #Semantics #eLearning

Archives

[November 2012](#)

January 2012						
S	M	T	W	T	F	S
	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30	31			

The Use of Metadata in the Description of e-Learning Content for Organic Agriculture

Currently there is a number of EU-funded projects that include the development of an agriculture-related training curriculum as one of their core activities, as well as an additional number of such projects that have already developed related content collections. However, in most cases, this content is not easily available through any mean of access, such as a portal, a course management platform or a website. This paper presents the development of the Organic Italian training curriculum, describing the use of metadata for the annotation of the training material through the use of a course management platform. It also provides an overview on the issues that were met during the metadata annotation process, including the manual translation of these metadata in the languages of the project.

Contributed by Maria Toader, Gheorghe Valentin Roman, Vanelia Prodanariu

You can view/ download the submitted paper from [HERE](#).

After reviewing the paper you can [leave your comment](#) on this post, and/or rate the paper (Rate this - above the title of the post).

Share this: [Leave a Comment](#)

Like this: [Like](#) Be the first to like this.

Filed under [Uncategorized](#)

[Leave a Comment](#)

Figura 95. Sito dell'esperimento VOA3R "Controlled Experiment for Reviewing Lifecycle Services": processo alternativo di revisione dei contenuti di ricerca in rete

L'obiettivo dell'esperimento è stato quello di coinvolgere tutte le parti interessate a contribuire allo sviluppo del processo alternativo di revisione dei contenuti di ricerca in rete. A tal scopo nel Blog dell'esperimento sono state caricate le versioni *full-text* di *paper* proposti al quinto Workshop Internazionale "Metadata and Semantics for Agriculture, Food and Environment" (MTSR 2012) [V. sito], i quali possono essere consultati, revisionati e commentati nello spazio apposito "Leave a comment". Al processo di revisione di ogni articolo presentato nel Blog possono unirsi diversi utenti registrati; tale approccio è proprio delle funzionalità VOA3R ed complementare al processo tradizionale e anonimo *peer review* cui risultati non sono disponibili apertamente [Šimek et al., 2012].