• Methodologies for cataloguing cultural heritage

• Computerized cataloguing and multimedia documentation

Spreading open data: use and reuse.

ICCD ongoing projects
Open data
Open data: definitions

Open data is the idea that some data should be freely available to everyone to use and republish as they wish, without restrictions from copyright, patents or other mechanisms of control.

It’s one of the main concept of Open government.
What is open government?

New concept of administration based on models, instruments and technologies allowing administration to be “open” and “transparent” towards citizens, to grant:

1. **public control over decisional process** of PA (TRANSPARENCY)
2. **participation to decisional process** of PA, trough bidirectional and shared interactions (PARTICIPATION and COOPERATION)

The instruments to achieve those goals are digital technologies

Relation between ADMINISTRATION and CITIZEN is based on TRUSTH
Transparency allows

- **public protection** *(right to information, integrity and impartiality of the administration)*,

- **improvement and simplification of the bureaucratic procedures** *(services’ quality)*
Open government in practice

USA, 2009

Memorandum for the Heads of Executive Departments on Transparency

“My Administration is committed to create an unprecedented level of openness in Government. We will work together to ensure the public trust and establish a system of transparency, public participation, and collaboration. Openness will strengthen our democracy and promote efficiency and effectiveness in Government.”

Art. 11 legislative decree 150/2009

«Transparency means full availability of information about every aspect of organization, even throughout publication on institutional websites of public administrations, (...) with the aim of promoting common forms of control, according to principles of good performance and impartiality.»

ITALY, 2009
Transparency and access

Transparency of a public administration is linked to the **free access of administrative data and information** by citizens, and to share documents and knowledges between institutions and local communities.
Data and rights
Spreading open data: use and reuse

Features of open data

Three meanings of open
1. They are published under licenses allowing their reuse by anyone, **also for commercial purpose** (**open license**);
2. in a format that can be easily processed by a computer and (**easily extracted**)
3. They are published in **formats defined by a published specification**, and which don’t require a proprietary software for their access (**open format**)

Examples:
Microsoft Excel is a format where data can be easily extracted, but it’s not an open format.
PDF is an open format but it’s not possible to extract data easily
CSV (comma separate value) is an open format and data are easily reusable (as XML format, etc.)
Features of open data

- **Complete**: every public data is available, except for data subjected to restrictions for privacy and security.
- **Primary**: data is collected from the source, with the maximum level of detail, not subjected to processing or manipulation.
- **Prompt**: data is publicly available as soon as possible, to preserve its value.
- **Accessible**: data is available to the widest range of users and for the widest types of purposes.
- **Machine-readable**: data is structured with the aim of being automatically processed.
- **Not biased**: data is available to anyone, without the necessity of registration.
Features of open data

- **Non-properitary**: data is available in format on which no one has an exclusive control.
- **Free**: use and reuse of data is not subjected to any copyright or patents’ restriction.
- **Reusable**: users can reuse and integrate data, even to create new resources, applications, programs and services for the community.
- **Researchable**: users can easily search data and information through search tools as databases, catalogues and search engine
- **Permanent**: all the described features are permanent during the entire lifecycle of data in the web
Spreading open data: use and reuse

Open data = common good and human rights?

• Data belongs to human kind (medical data, environmental data, meteorological data etc.)
• Data of public administration is funded by public money
• Knowledge of data allows the growth of society (development of app reusing open data)
• Data is essential to ease common human activities (for example cartography, etc.)
• In scientific field discovery rate grows thanks to the data access
Public Sector Information (PSI)

Data of public administrations that are non personal or anonymous, produced by a public institution within its mandate and that, if open and available via web, can increase transparency and promote collaborative interaction between citizens and PA.

Data is the ‘money’ in the knowledge society, where richness is measurable in function of range and speediness of exchange and reuse of data between datasets.

Open data paradigm stimulates PA to recover and organize information heritage created during the year and often underused, because of dispersion into lot of offices and/or archival format.
Regulatory framework

https://www.dati.gov.it/content/riferimenti-normativi-documenti-indirizzo
Art. 2 it – definition of publication:
Access
-in institutional website of public administrations
-to documents, information and data about organization and activities of public administrations.
-For everyone, without authentication and identification

Decreto trasparenza – d.lgs 33/2013
Spreading open data: use and reuse

Museums, Archives and Libraries

Directive 2013/37/UE on reuse of information in the public sector (c.d. PSI - Public Sector Information) *

Main elements:

a) **application** of the directive to libraries, museums and archives;

b) obligation to allows reuse of PSI for commercial and non commercial purposes became a general principle for each Member State

Acknowledged in Italy with DECRETO LEGISLATIVO 18 maggio 2015, n. 102
(«Attuazione della direttiva 2013/37/UE che modifica la direttiva 2003/98/CE, relativa al riutilizzo dell'informazione del settore pubblico»)
Guidelines for publishing Open Data by PPAA

Part I – PA and Open data
It introduces the concept of Open Government, the practice of Open Data and it points out the related regulatory framework.

Part II – How to open data of PA
It describes technical, organizational and juridical aspects to analyze, before publishing Administration data.

http://egov.formez.it/sites/all/files/VademecumOpenData_0.pdf
CAD - Codice dell’Amministrazione Digitale

(Code for Digital Administration)

Dispositions about the use of digital technologies as main tool in interactions between public administration and citizens.

Chapter V: data of public administrations and on-line services
CAD - Availability to privates

Art. 50 – Availability of data in public administrations.

• “PA data are formed, gathered, stored, and **made available**, through the use of digital and communication technologies, which allows **their reuse and access by public administrations and privates**, according to conditions stated by the law”

• exceptions:
  - restrictions stated by laws and rules,
  - restrictions in terms of personal data protection
  - restrictions stated by European Union in terms of reuse of information in public sectors.
Art. 52 – On-line access and reuse of data of public administrations.

Data and documents published by public administrations in any way, without the adoption of a specific license, according to art. 2, item 1, letter h), of decreto legislativo 36/2006, are meant to be released as open data, according to art. 68, item 3, of the Code. The possible adoption of a specific license must be justified according to national guidelines in item 7.

- In defining contracts related to products and services implying gathering and management of public data, public administrations provide for clause to allow on-line access and reuse, by physical and juridical persons, of those data, of their metadata, data structures and databases.
• Art. 52 – On-line access and reuse of data of public administrations.

Agency for digital Italy promotes valorization of national informative heritage and carries out dispositions according to Chapter V of the Code.

The Agency sends to Presidente del Consiglio dei Ministri or to Minister responsible for technology innovation
- the National Agenda to define contents and goals of valorization interventions
- an annual report on the status of the process of valorization in Italy; this report is published in open format on institutional website of Presidenza del Consiglio dei Ministri.
Problems and considerations
Real problems 1/2

- Organizational problems:
  - Generally Public Administration is not fully informed of considerable heritage of data it has;
  - Only a little part of PA data is available in digital format and, if it is, interoperability is not always possible;
  - Often licenses don’t allow republishing and reuse.

Source: Ernesto Belisario, lawyer and expert in law about new technology
Real problems 2/2

- Increase the awareness in citizens that their request of transparency corresponds to an higher availability of open data, changing anger and frustration into civil activism;
- Promote action of media information literacy
- Promote, in information technology sector, growth and education of qualified figures for example in data mining, in management and use of open data, in web app development for the access to open data
False problems

- What if people interpret data in a wrong way? *(What if people vote in a wrong way?)*
- What if data are used by few people? *(it is always a sign of transparency)*
- What if my data are wrong? *(publish them online and ask community for help)*
- What if data prove my inefficiency? *(inefficiency is a problem independent from the publication of data. Be aware of the problem and solve it)*

Source: Ernesto Belisario, lawyer and expert in law about new technology
Privacy problems?

The Guarantor for protection of personal data has showed his approval to Open Data:

*Deliberation n. 88/2011*

Achieving transparency of PPAA activities is possible also without using personal data. “*It’s not necessary to adopt specific caution, if public administrations publish on web site information not ascribable to identified or identifiable persons*”
Data can be considered as a **product of an intellectual elaboration/creation**

→ Every work, result of an intellectual work of an author, is **protected** from a moral and patrimonial point of view.

To allow **free access to public data**, one of the main tools is the license.

We can distinguish:

- **Open license**
- **Closed license**.
Copyright problems

Closed license - **copyright:**
- License agreement allowing **only the author** to exercise the **rights to the use and manipulation** of the work (exclusive)

Open license - **copyleft:**
- License agreement containing a specific **list of permitted uses for the work** (permissive)
- In general a work can be copied and distributed without costs
Example of open license

Creative Commons Licenses

You can choose what right you want to left and what right you want to maintain, by combining four terms of distribution:

1. **BY** - Attribution (obligation to give the author the credits)
2. **NC** - Non-commercial (obligation to copy, distribute the work and make derivative works based on it only for non-commercial purposes)
3. **ND** - Non-derived (any derivative work is permitted)
4. **SA** - Share-alike (obligation to distribute derivative works only under an identical license)
Spreading open data: use and reuse

Types of license

Information about the type of license is an essential metadata for the reuse of dataset. It must be always indicated, showing the name, the version and the reference to the text of the license.

Every license not allowing derivative works or commercial purposes (i.e., CC license with NC and ND terms) and/or stating terms that restrict uses and distributions of data, isn’t valid to identifying open dataset.

http://lg-patrimonio-pubblico.readthedocs.io/it/latest/licenzecosti.html
Let’s make data free!
Data suppliers stimulate the creation of market

Forward-thinking public administrations have understand the importance, even for themselves, of promoting the development of applications and the mashup with open data they have and spread

USA and UK government and the World Bank itself, in their websites have a section for “apps”
Example for the reuse of open data

Berners-Lee in the TED-2010 map created by a lawyer to prove the correlation between houses where white people live and houses connected to aqueduct, showing the racial discrimination against black people in Zanesville (Ohio, USA). Thanks to this map the lawyer convinced the judge to sentence the county to a compensation of over 10 million dollars.
WHERE DOES MY MONEY GO?

Showing you where your taxes get spent

122.4B

Environment

170.5B

Helping Others

501.9B

GRAND TOTAL (2010)

Our Streets

Running Government

Culture

Order & Safety

Running The Country, Social Systems

Expenditure on
Grand Total (2010)

17.7B - 23.5B
23.5B - 27.5B
27.5B - 31.8B
31.8B - 43.3B
43.3B - 47B
47B - 105.3B (£)
### Civil hacking? Open Camera (2010)

I conti segreti della Camera dei Deputati resi "liberi"

<table>
<thead>
<tr>
<th>Tipo</th>
<th>Settore</th>
<th>Fornitore</th>
<th>Oggetto</th>
<th>Previsione 2010</th>
<th>Note</th>
<th>Ordinato al 1/2/2010</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contratti</td>
<td>Varie</td>
<td>CEDAT 85 SRL</td>
<td>SERVIZIO DI SUPPORTO ALLE ATTIVITA' DI RESOCONTAZIONE STENOGRFICA</td>
<td>66,097.77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contratti</td>
<td>Varie</td>
<td>TECNOCONFERENCE EUROPE SRL</td>
<td>NOLEGGIO IMPIANTI TRADUZIONE</td>
<td>130,000.00</td>
<td>gara in corso</td>
<td></td>
</tr>
<tr>
<td>Contratti</td>
<td>Varie</td>
<td>CEDAT 85 SRL [PROGETTO LAVORO SOC. COOP.]</td>
<td>ASSISTENZE OPERATIVE A SUPPORTO DELLA GESTIONE DOCUMENTALE E TECNICA</td>
<td>3,135,000.00</td>
<td>da gara</td>
<td>2,869,548.00</td>
</tr>
<tr>
<td>Contratti</td>
<td>Varie</td>
<td>INA ASSITALIA (rti con FONDIARIA, UGF)</td>
<td>CONVENZIONE ASSICURATIVA</td>
<td>2,454,050.34</td>
<td>da gara/gara in corso</td>
<td></td>
</tr>
<tr>
<td>Contratti</td>
<td>Telecomunicazioni</td>
<td>RAI RADIOTELEVISIONE ITALIANA SPA</td>
<td>CANONE DI ABBONAMENTO ALLA</td>
<td>3,905.90</td>
<td>fuori campo applicazione</td>
<td></td>
</tr>
</tbody>
</table>
Esempio: la mappa della salute

consente di esplorare diversi indicatori relativi alla salute della popolazione del Galles, tra cui le degenze ospedaliere (costo, durata, tempi di attesa), l’incidenza di particolari malattie, le cause e i tassi di mortalità, le maternità e la salute dell’infanzia.

Esempio: i dati relativi alla mortalità per abuso di alcool sulla mappa.
Spreading open data: use and reuse

Esempio

STREETLINE
Connecting the Real World

PARKSIGHT™ - FOR CITIES

From street-level sensors to mobile apps and parking analytics, Streetline's technologies help you truly understand and effectively manage parking in your city. Learn More »

PARKEDGE™ - FOR GARAGES

Publish your parking inventory to Parker™, the leading parking guidance app. Increase occupancy. Accept reservations. Drive revenue goals. Learn More »

PARKER™ - FOR MOTORISTS

Everything you need to know about parking in one app. With real-time parking availability via Parker, quickly and easily find, pay for, or reserve available parking. Learn More »

Computerized cataloguing and multimedia documentation
Dati aperti per migliorare la qualità della vita
Are You Safe lets you know how safe you are at all times based on your current location within the city. All it takes is a quick tap to see a threat meter of your safety level along with hyperlocal crime data broken down by type.

Features:
- Simple visual representation of your safety level on the Threat Meter
- Up to date crime data from police and local city governments
- Dynamically updated recent crime data for your surrounding area
- Walk/drive/bike/bus/canoe around and see your status update as you go

Uses:
- Visiting and unfamiliar with the city?
- Debating whether to walk or take a cab?
- Headed to an area you haven’t been to before?
- Not sure if you should park your car on the street?

All Cities:
- Atlanta, Washington D.C., Sacramento, Indianapolis, Milwaukee, Chicago, San Francisco, Dallas, Hampton Roads (VA Beach, Norfolk, Chesapeake, Newport News, Suffolk, Portsmouth), Cincinnati

Support:
- For new features and updates, follow us on twitter
- Drop us a line for help or media inquiries

Disclaimer: AreYouSafe uses actual crime statistics to derive safety ratings. This data is sourced from police departments and local city governments. AreYouSafe makes no claims as to the accuracy or validity of this data. The application is for novelty purposes only. Decisions should not be made based on the information yielded by the.

Computerized cataloguing and multimedia documentation

Ministero dei beni e delle attività culturali
Istituto centrale per il catalogo e la documentazione

Ministry for Cultural Heritage and Activities
Central institute for cataloguing and documentation
MiaPA utilizza la rubrica della PA realizzata da Formez PA per geolocalizzare gli uffici della pubblica amministrazione
Spreading open data: use and reuse

Venezia News

Computerized cataloguing and multimedia documentation
Linked Open Data
Spreading open data: use and reuse

I Linked Open Data

Pubblicate le “Linee guida per l’interoperabilità semantica attraverso i linked open data”

La Commissione di Coordinamento SPC ha approvato le “Linee guida per l’interoperabilità semantica attraverso i Linked Open Data (PDF)”, consultabili anche nella sezione dedicata alle attività della Commissione di...

2012
The main feature of LOD is its double value:

1. “open” → to increase transparency and reuse by anyone
2. “linked” → based on the main technologies of web of data

Data that is linked to other data referring to identical objects or to objects that are related

Availability of Open data is important to increase information, but openness is not enough. Open data has to be autodescriprive to allow inferences thanks to aggregation and correlation of datasets. Semantic web technologies, and in particular Linked Open data model, are important tools to overcome Open data restrictions.
WEB (World wide web) is born in 1991, when Tim Berners Lee launched the first web page. This first stage of web is called web 1.0:
- read-only (passive users and unidirectional information flow)
- Three main elements - URI ((Uniform Resource Identifier - system of identification of a web resource)
  - HTTP (protocol for the transmission of information, independently from the type of data)
  - HTML (markup language to make web document readable by individuals)
- Problems: - information is fragmented (formats, databases, not-explicit relations)
  - Difficulty in information retrieval (caused by information overload)
The following stage of web is called **web 2.0** (or **semantic web**) evolution of Web, toward its transformation into a **unique, huge database**, linking **datasets** to find out new information that can be **freely republished**
The **semantic interoperability** is the ability of computer systems to exchange data with **unambiguous, shared meaning**.

And the **semantic web** is the place where semantic interoperability can be developed: the future of the web is linked to semantic technologies, to realize a new **machine readable** and **understandable** web,

**a web of data linked through semantic relation, processable by machines**

Only with technologies linked to SW it will be possible to show data about cultural heritage **linked with a universe of data**, on which machines will be able to make **automatic reasoning**, thanks to their huge computing power.
Developed by Berners-Lee to explain the various protocols and challenges underlying semantic Web technologies.
**Unicode** is a system to **codify characters**. It associates each character of every language to a unique number, that is the same for every computer system, software or language used.

The **Uniform Resource Identifier (URI)** allows to name every object in a **non-ambiguous** way

https://it.wikipedia.org/wiki/Michelangelo_Buonarroti

(if we use HTTP protocol we can access it via WEB).
XML + Namespace + XML Schema allows to identify structure and syntax of web documents, through the representation of textual contents that are hierarchically organized.

**WEB Semantic pyramid**

**XML (eXtensible Markup Language):** standard defining the syntax to mark up data through tags

```xml
<title>Proud and preguidice</title>
<author> <firstName>Jane</firstName><lastName>Austen</lastName> </author>
```

- **non-proprietary textual format** (it can be used by every computer system)
- Based on association of descriptive tag to data
- We can choose any name for the tag
Namespace: it’s the domain, identified by a URI and a prefix, that allows to solve conflict between names of tags.

Example:
In the domain of books, the tag `<title>` means the name of a book.
In the domain of ‘persons’, the tag `<title>` means someone’s position or job.

To avoid conflict, we can identify the two domains as follow:

Book domain: URI www.books.it
prefix book
tag `<book:title>`

Person domain: URI www.persons.it
prefix person
tag `<person:title>`
WEB Semantic pyramid

- Superior levels are still “work in progress”

- RDF + RDF Schema allow to express machine-processable statements, throughout triples having the form of Subject, Verb and Direct Complement. RDF Schema adds to RDF some concepts to increase expressive potentiality of RDF.

- Ontology and vocabulary make available instruments to define main concepts to “speak” about a certain domain.
RDF: a new language for the web

- **RDF-Resource Description Framework** is the grammar of the language through which the web evolution is possible.

- **RDF** is a **model to represent relations between data**, so that they can be understood not only by humans, but also by computer.

According to W3C, **RDF** is important to the evolution of web from *machine-representable* to *machine-understandable*. The idea is to generate “documents” that can be read and understood by human beings but that also can be interpreted by automatic systems.
**RDF: a new language for the web**

**Prerequisite: granularity of data**

“it refers to the **size** in which data fields are subdivided”.

**Example: Postal address**

**Low granularity:**
via di San Michele 18, 00153.
Roma (RM), Italy

**Fine granularity:**
Street address: Via di San Michele
Address number: 18
Postal code: 00153
City: Roma
Province: (RM)
Country: Italy
RDF is inspired to logic of predicates: data is expressed through simple statements, made up by triples in the form of “subject-predicate-object”.

Every element of the triple is an atomic, meaningful concept, identified by a URI:
- Subject: resource, identified by a dereferenciable URI (URI allowing the access to the description of a resource – HTTP URI)
- Predicate: property of the resource, identified by a URI (every domain can define property of resources thanks to ontologies)
- Object: resource or literal
  Not identified by a URI, but by a data type. For example the object can be a date of birth, that is not identified by an URI, but by the digit model yyyy/mm/dd
RDF in practice

Computerized cataloguing and multimedia documentation
<table>
<thead>
<tr>
<th>SOGGETTO</th>
<th>TIPO_SCHEMA</th>
<th>CODICE_UN</th>
<th>LOCALIZZAZIONE</th>
<th>CONTENITORE</th>
<th>DATAZIONE</th>
<th>AMBITO_CU</th>
<th>AUTORE</th>
</tr>
</thead>
<tbody>
<tr>
<td>David</td>
<td>DA_3.00</td>
<td>09 00281988</td>
<td>Italia, Toscana, FI, Firenze via Ricasoli</td>
<td>monastero, benedettino ferro</td>
<td>&quot;sec. X VI 1501&quot;</td>
<td>Buonarroti Michelangelo</td>
<td></td>
</tr>
</tbody>
</table>

Granularity

Every line of the table is a resource.
Columns are properties.

Resource identifier

Resource name

Other features/properties of resource
Spreading open data: use and reuse

URI importance

Data describes things, persons, places, books, artifacts, institutions, companies, etc.

Those things have identifying names (Dante, Palermo, David di Michelangelo). But in RDF, those names are only labels and the URI is the real name identifying the resource.

URI identifies unambiguously a resource.
A resource can be described in more than a triple and have different functions.

Triples can share object or subject, to form a graph.
4 rules for Linked Open Data

Tim Berners-Lee in 2006 identified four rules to publish Linked Open Data:

1. To use URI (Uniform Resource Identifier) as names of resources
2. To use HTTP URI (that is URL), so that it is possible to access the resource via web.
3. To use open standards to describe resources (RDF and SPARQL).
4. To include link to other URI, to promote the discovery of other resources
Open data’s five stars

**LINKED DATA**

- ★ On the web, open license
- ★★ Machine-readable data
- ★★★ Non-proprietary format
- ★★★★ RDF standards
- ★★★★★ Linked RDF

IS YOUR DATA 5★?
Spreading open data: use and reuse

Our triple in the web of data

restauro

2004

ha data

ha autore

ha evento

ha bibliografia specifica

ha bibliografia specifica

Vasari G.
1962
00001614
pp. 19-23

Condovi A.
1823
00001558
pp. 21-23

1501-1504

ha data creazione

ha data creazione

creato da

proprietà di

Ministero
dei beni e delle
attività culturali
e del turismo

Gallerie dell’Accademia

Firenze

Conservato a

Our triple in the web of data

Computerized cataloguing and multimedia documentation
Spreading open data: use and reuse

---

**Computerized cataloguing and multimedia documentation**

Ministero dei beni e delle attività culturali
Istituto centrale per il catalogo e la documentazione

---

**David**

Soggetto di: Gallerie dell'accademia

Sito a: Firenze

---

**Michelangelo Buonarroti**

Enciclopedia dei ragazzi (2006)

Buonarroti, Michelangelo
Dizionario Biografico degli italiani (1972)

MICHELANGELO BUONARROTI
Enciclopedia Italiana (1934)

---

**Luce 90**

La settimana di Roma 1948
7/7/1948
Messaggio dello stato italiano: il "Deibel" di Michelangelo, partito per l’America.

Estratto da la settimana di Roma: "le opere di Michelangelo sono allora disposte a Firenze e alla Torre di Babele."

---

**OPAC SBN**

Catalogo del servizio bibliotecario nazionale

---

**Central institute for cataloguing And documentation**
Spreading open data: use and reuse

Linked Open Data Cloud
May 2007

http://linkeddata.org
Spreading open data: use and reuse

Linked Open Data Cloud
September 2011
Computerized cataloguing and multimedia documentation

Ministero dei beni e delle attività culturali
Istituto centrale per il catalogo e la documentazione

Linked Open Data Cloud
Agosto 2017
Main experimentations with linked open data in the cultural heritage sector
Europeana.eu, Europe’s cultural heritage portal

26M objects from 2,200 European galleries, museums, archives and libraries
In 1974 FAO set up an initiative called AGRIS to make information on agriculture research globally available — now 4 million of bibliographic records.

AGRIS bibliographic records sometimes suffered of lack of information, like the full text of a document → only 3% of the entire collection had a full text link, so the user had to search Google to find the publication.

Users, and especially researchers, are interested in the entire content of a publication and not only in its abstract. Moreover, lack of information like connections with related work and related web resources makes impossible to interlink to other sources of data.

To allow users to access the fulltext of a publication and all the information the Web knows about a specific research area in the agricultural domain, OpenAGRIS has been developed, following Linked Open Data principles.

http://agris.fao.org/openagris/
Spreading open data: use and reuse

Computerized cataloguing and multimedia documentation
Spreading open data: use and reuse

Data from Europeana

Rice irrigation in Texas...

Taylor, Thomas Ulvan

Language: English

Data provider: Library of Congress (archive.org)

Type: TEXT

Go to reference

The greatest of grains. Rice.

On the growth of the rice plant, by Shinkichi Suzuki.

The storage and germination of wild rice seed.

Gibberella fujikuroi (Sawada) Wollenweber, the new parasitical fungus on rice in the Republic of Macedonia

Computerized cataloguing and multimedia documentation

Ministry for Cultural Heritage and Activities
Central institute for cataloguing and documentation
Spreading open data: use and reuse

Title
Rice irrigation in Texas...

Related Titles
Series: Bulletin of the University of Texas, no. 16

By
Taylor, Thomas Ulvan.

Genre
Book

Publication info
Austin, Tex., Von Boeckmann, Schutze & co., state printers, 1902.

Subjects
Irrigation, Rice, Texas

DOI
http://dx.doi.org/10.5962/bhl.title.35260

Find in a local library

View Book

Copyright & Usage:
Not specified

Download book: PDF, ORC, JPG, ALL

Contributing Library:
Library of Congress

Sponsor:
Sloan Foundation

Date Scanned:
07/23/2009
DBpedia at the center of linked data cloud

→ DBPedia is one of the main nodes of the current linked data cloud

→ Using its URI is considered as authoritative reference
Ontologies:
ArCo
Defining ontology

“a specification of a conceptualization”
T. R. Gruber

A conceptualization is a formal representation of a cluster of knowledges: it is a cluster of objects, concepts and relations between them belonging to a particular area of study (domain).
A conceptualization is an abstract and simplified representation of a particular domain that we want to represent for any purpose.
### Defining ontology

<table>
<thead>
<tr>
<th>An ontology has a common language</th>
<th>syntax</th>
</tr>
</thead>
<tbody>
<tr>
<td>The meaning of the symbols and expressions in an ontology is clear</td>
<td>semantics</td>
</tr>
<tr>
<td>Symbols and expressions with similar semantics are grouped in classes</td>
<td>conceptualization</td>
</tr>
<tr>
<td>Concepts are organized in a hierarchical way</td>
<td>taxonomy</td>
</tr>
<tr>
<td>Implicit knowledge can be made explicit</td>
<td>reasoning</td>
</tr>
</tbody>
</table>
An ontology is a **formal model** representing a **knowledge domain**, according to specific requirements. It is used to describe the **semantic of data** with an **established terminology** and it can be **reused** in other projects with similar goals.

To define an ontology we need:

1. To collect requirements
2. To **define classes**
3. To **organize** classes in a taxonomic **hierarchy** (subclasses-superclasses)
4. To define **properties and describe restrictions** for each of them
The agreement with CNR

Started from December 2014, the agreement defines cooperation between MIBAC and CNR (Laboratorio di Tecnologie Semantiche or STLab) to:

- Model place of culture/events data according to paradigm of Linked Open Data
- Find out technical solutions for the integration and rationalization of databases of the cultural heritage (promotion and protection)
- Cooperate in spreading the reuse of open data and in defining guidelines for the promotion of cultural heritage
MIBAC manages “Places of culture” and “Cultural events databases”, containing information on:

- About 8000 places (archaeological areas and parks, monuments, monumental complex, other permanent structures for the public access of heritage)
- Cultural events (exhibitions, conferences, seminars, catalogue presentations, ecc.) organized by the Minister and linked institutions

Database features:
- **open-data** format (XML),
- **data structure** available in [http://www.beniculturali.it/mibac/xsd/MibacSchema-1.2.xsd](http://www.beniculturali.it/mibac/xsd/MibacSchema-1.2.xsd)
- **open license** CC-BY 3.0
## Datasets

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Luoghi della cultura statali</td>
<td>Estratto dal database dei luoghi della cultura (DBUnico 2.0) relativo ai luoghi della cultura statali.</td>
</tr>
<tr>
<td>Contenitori fisici</td>
<td>Dataset relativo alle schede ICCD per i contenitori fisici, entità che rappresentano “nodi di aggregazione” di beni culturali mobili e che indicano il luogo fisico (un edificio, un complesso architettonico o uno spazio territoriale) dove è collocato un bene.</td>
</tr>
<tr>
<td>Anagrafe delle biblioteche italiane</td>
<td>Dataset dell’Anagrafe delle biblioteche italiane, rilasciato dall’Istituto Centrale per il Catalogo Unico (ICCU)</td>
</tr>
<tr>
<td>Archivi di Stato</td>
<td>Dataset dell’Anagrafe degli Archivi di Stato, rilasciato dall’Istituto Centrale per gli Archivi (ICAR)</td>
</tr>
<tr>
<td>Archivio schede di catalogo</td>
<td>E’ l’inventario “topografico” dell’archivio cartaceo delle “vecchie” schede di catalogo utilizzate per il rilevamento dei beni immobili architettonici ed archeologici e per i beni mobili artistici, storici ed archeologici, pervenuto all’allora l’Ufficio Centrale per il Catalogo nel 1969. Dataset rilasciato dall’Istituto Centrale del Catalogo e della Documentazione (ICCD)</td>
</tr>
<tr>
<td>Fondo MPI</td>
<td>Set minimo di dati dell’Archivio fotografico della Direzione generale antichità e belle arti del Ministero della pubblica istruzione, acquisito dall’ICCD. Dataset di test rilasciato dall’Istituto Centrale del Catalogo e della Documentazione (ICCD)</td>
</tr>
</tbody>
</table>
Aim: linked (MIBAC) data

- Photographs (ICCD)
- Places of culture / «containers»
- Catalogue sheets (ICC D)
- Archives (SAN-ICAR)
- «Touristic information»
- Books (SBN – ICCU)
Result: Cultural-ON

✓ **Multilingual** ontology, not to limit its reuse to national boundary
  • Available in English and Italian

✓ Definition of classes and properties to point out **essential elements** of Institutes and Places of Culture, as they are defined by the Codice dei Beni Culturali

✓ **In compliance** with recommendations of **AgID guidelines** about semantic interoperability through Linked Open Data
Main elements of the Ontology

Main class, modeling the concept of **Cultural Institute or Site (CIS)**
Spreading open data: use and reuse

**Site**

<table>
<thead>
<tr>
<th>Property</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>hasAddress</td>
<td>Address</td>
</tr>
<tr>
<td>hasContactPoint</td>
<td>ContactPoint</td>
</tr>
<tr>
<td>hasGeographicalLocation</td>
<td>GeographicalFeature</td>
</tr>
<tr>
<td>hasGeometry</td>
<td>Geometry</td>
</tr>
<tr>
<td>hasNameInTime</td>
<td>NameInTime</td>
</tr>
<tr>
<td>hasSiteDescription</td>
<td>SiteDescription</td>
</tr>
<tr>
<td>hosts</td>
<td>Event</td>
</tr>
<tr>
<td>isSiteOf</td>
<td>CulturalInstituteOrSite</td>
</tr>
<tr>
<td>isSubjectOf</td>
<td>CreativeWork</td>
</tr>
<tr>
<td>description</td>
<td>Literal</td>
</tr>
<tr>
<td>name</td>
<td>Literal</td>
</tr>
</tbody>
</table>

**Location of a place or an event**

Computerized cataloguing and multimedia documentation

Ministero dei beni e delle attività culturali
Ministry for Cultural Heritage and Activities
Central institute for cataloguing and documentation
The name of the place can **vary** during the **time**.
Access restrictions

✓ **Information** about place **access**

- E.g., type of booking, temporary closure, opening hours, type of access, statistics on

![Diagram](image-url)
Spreading open data: use and reuse

Contact point and services

**Contact point** of the place/headquarters of cultural place/event

<table>
<thead>
<tr>
<th>ContactPoint</th>
</tr>
</thead>
<tbody>
<tr>
<td>available : OpeningHoursSpecification</td>
</tr>
<tr>
<td>hasCertifiedEmail : anyURI</td>
</tr>
<tr>
<td>hasEmail : anyURI</td>
</tr>
<tr>
<td>hasFax : anyURI</td>
</tr>
<tr>
<td>hasTelephone : anyURI</td>
</tr>
<tr>
<td>hasWebsite : anyURI</td>
</tr>
</tbody>
</table>

**Services offered** by the cultural place

Computerized cataloguing and multimedia documentation
Spreading open data: use and reuse

**Tickets**

<table>
<thead>
<tr>
<th>Ticket</th>
</tr>
</thead>
<tbody>
<tr>
<td>forAccessTo : CulturalInstituteOrSite or Event</td>
</tr>
<tr>
<td>hasValidity : Validity</td>
</tr>
<tr>
<td>description : Literal</td>
</tr>
<tr>
<td>name : Literal</td>
</tr>
</tbody>
</table>

**Offers**

ha come super-classi

- include\(^{op}\) some Biglietto\(^{c}\)
- ha validità\(^{op}\) only Validità\(^{c}\)
- descrizione\(^{dp}\) only literal

è nel dominio di

- ha specifica di prezzo\(^{op}\), ha utente elegibile\(^{op}\), include\(^{op}\)

è nel codominio di

- offre\(^{op}\)
Management of normal opening hours, used for Events, Places of Cultures and for their contact point

We can manage also special opening hours through the property description