

AthenaPlus workshop
**Innovative tools and pilots for access to digital cultural heritage in
the framework of Europeana and national systems**

*Experiments of Linked Open Data in the
framework of the National Union
Catalogue (SBN)*

Patrizia Martini (ICCU)
Achille Felicetti (VAST-LAB – PIN)

The Central Institute for the Union Catalogue of Italian Libraries (ICCU)

- ❑ Is the national center which promotes and coordinates **cataloguing** and **digitisation** activities of the Italian libraries
- ❑ Is responsible for guiding, producing, adapting and disseminating the **standards and rules for cataloguing and digitisation** of the library heritage
- ❑ Is in charge to coordinate and to manage important national systems and census: National Library Service (SBN), National Census of Italian XVI century editions (EDIT16), Census of Manuscripts of Italian Libraries (Manus), Internet Culturale, CulturalItalia...
- ❑ Is in charge to coordinate many European and international projects in the field of digital cultural content

<http://www.iccu.sbn.it>

ICCU institutional tasks

- ❑ Promotion and production of cataloguing rules and guidelines and diffusion of international standards
 - SBN Cataloguing rules
 - Translation of IFLA international standards and models (ISBD, FRBR...)
 - Translation of RDA rules (LoC, JSC)
- ❑ Activities in close connections with international bodies: IFLA, PUC, VIAF, EURIG, Dublin Core Metadata Initiative...

Experiments of LOD in SBN

The SBN National Coordination Committee has committed to ICCU a study for the definition of a data model for publishing OPAC SBN data in Linked Open Data

Initial steps:

- ☐ Analysis of a basic methodology for creating open linked data from the bibliographic data according to the international standards and cataloguing rules in SBN
- ☐ Drawing the logic scheme relating to bibliographic content, attribution of descriptive, semantic, management metadata
- ☐ Selection of a first set of bibliographic records and export them from OPAC SBN in UNIMARC format
- ☐ Development of namespaces to accommodate UNIMARC element sets and value vocabularies following the FRBR family of models

The working group:

Francesco Gandolfi, Patrizia Martini, Elisa Sciotti (ICCU), Achille Felicetti (PIN)

Experiments of LOD in SBN

Further steps:

- Data conversion into RDF(Resource Description Framework) encoding syntactic basis of the Semantic Web and linked data.
- Growing an interface that allows libraries to connect to the SBN Index, and derive triples that constitute entities of their interest (work, expression, etc.), converted to standard formats, and afterwards usable within their catalogue
- Creation of links to data sets of international open linked data projects: VIAF, Europeana, Linked Heritage, DBPedia, etc.
- Publication of data in RDF/XML with different modes: with appropriate access URI and URL, through interfaces, SRU, OAI SPARQL Endpoint
- definition of EP-SPARQL search function for published datasets consultation
- validation, ensuring the respect of the formats required by the semantic web, and multiple tests on the data set functionality in a manner consistent with the principles and purposes of the semantic web



Overview

- ☐ Evolution of library Linked Data RDF Description Framework and the Semantic Web
- ☐ Identity management
- ☐ Schema, mapping, interoperability
- ☐ Open publishing
- ☐ Identifiers
- ☐ Universal Bibliographic Control
- ☐ Role of national libraries and bibliographic agencies
- ☐ Bibliographic standards (IFLA)
- ☐ Impact and implications for catalogues and cooperation

<http://www.ifla.org/>

<http://www.gordondunsire.com/presentations.htm>



□ IFLA 2013 “Expanding MARC Metadata Services with Linked Open Data”

This session focused on concepts, strategies, experiences, technical issues and prospects of linked open data (LOD) applied to MARC and UNIMARC records.

Linked data initiatives

- UDC Consortium: published a selection of around 2,000 UDC classes in 16 languages online as the UDC summary (RDF version in development)
- Virtual International Authority File (VIAF): a set of linked controlled vocabularies (authority records of personal names) by national bibliographic agencies
- ISBD: prescribes vocabulary control for the data in the Area 0 for content form and media type. Terms for the elements (content form, content qualification, and media type) are taken from closed lists

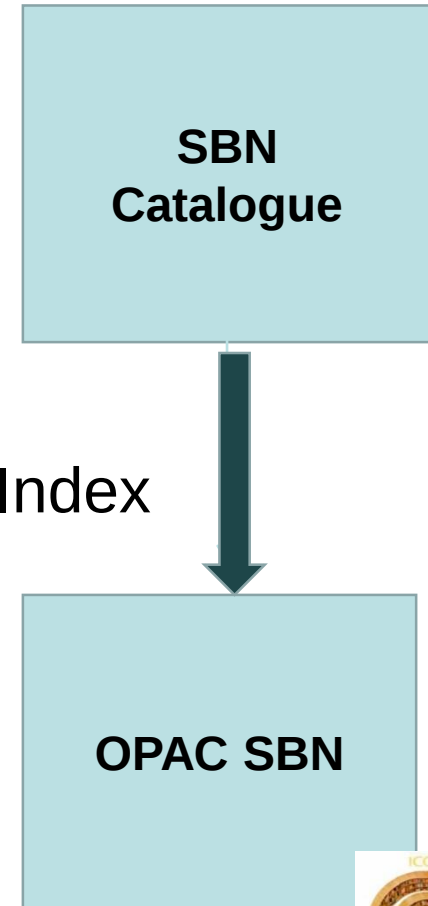


5.000 libraries linked to the SBN Catalogue through 84 nodes

- 12 M **resources**
- 3,5 M authors
- 66 M locations

Daily traffic

- 200.000 daily requests from libraries to Index
- 52 M searches



SBNMARC protocol

SBNMARC is a communication protocol that simplifies interacting between LMSs and Index, because it uses XML messages and a schema based on **UNIMARC semantic**: concepts and attributes.

Cataloguing in SBN

Shared cataloguing is based on:

- ☐ a «core» of common data
- ☐ specific data for different types of resources (Music, graphic, cartographic)
- ☐ authority files:
 - Uniform titles
 - Names (personal and corporate body names)
 - Printers' devices
 - Places of publication
 - Semantic headings (subjects, classification)

Modern: monographs and periodicals post-1830

Antique: monographs until 1830.

Music: printed and manuscript music, including music librettos, musical sound recording, musical videos

Graphic: drawings, paintings, posters, pictures, photographs, etc.

Cartographic: Manuscript and printing cartographic resources

esempio_sbn.jpg - Visualizzatore foto di Windows

File Stampa Posta elettronica Masterizza Apri

OPAC SBN
CATALOGO DEL SERVIZIO BIBLIOTECARIO NAZIONALE

ICCU Ministero del bene e della cultura e del turismo DGBIB

Cerca nel Catalogo

Home Informazioni Aiuto Glossario Contatti English Accessibilità

Ricerca: Identificativo SBN = IT\ICCU\TO0\1802857 (parole in AND)

Formato Etichette Nascondi biblioteche Stampa E-mail

← Lista sintetica

Scheda: 1/1 [Scheda Unimarc](#) Scarico Unimarc Export Endnote Export Refworks Citazioni [Aggiungi a preferiti](#)

Livello bibliografico Monografia
Tipo documento Testo a stampa
Autore principale Pirandello, Luigi <1867-1936>
Titolo **Racconti per una sera a teatro / Luigi Pirandello ; a cura di Guido Davico Bonino**
Pubblicazione Palermo : Sellerio, 2011
Descrizione fisica 410 p. ; 17 cm
Collezione **La memoria** ; 844
Numeri **[ISBN]** 88-389-2536-4
[ISBN] 978-88-389-2536-8
[BNI] 2011-882
Nomi **[Autore]** **Pirandello, Luigi <1867-1936>** scheda di autorità
Davico Bonino, Guido scheda di autorità
Classificazione Dewey **853.914 (21.)** NARRATIVA ITALIANA, 1945-1999
Lingua di pubblicazione ITALIANO
Paese di pubblicazione ITALIA
Codice identificativo IT\ICCU\TO0\1802857

Per accedere ai servizi di prestito interbibliotecario e fornitura riproduzioni ILL SBN selezionare una o più biblioteche e cliccare qui

Dove si trova

☐	AG0091	PALBA	Biblioteca Museo Luigi Pirandello - Agrigento - AG
☐	BA0018	BA18A	Biblioteca nazionale Saganiga Visconti-Volpi - Bari - BA
☐	BL0061	VIASE	BIBLIOTECA CIVICA DI SEDICO - Sedico - BL
☐	BO0239	UBODA	Biblioteca comunale Cesare Pavese - Casalecchio di Reno - BO
☐	CN0037	TO050	Biblioteca civica - Cuneo - CN
☐	CN0065	TO055	Biblioteca civica - Mondovì - CN
☐	CT0062	PALBC	Biblioteca regionale universitaria - Catania - CT
☐	FC0018	RAVFO	Biblioteca comunale Aurelio Saffi - Forlì - FC
☐	FI0098	CFICF	Biblioteca nazionale centrale - Firenze - FI
☐	FI0101	CFIMF	Biblioteca Marucelliana - Firenze - FI
☐	FR0021	RMSE1	Biblioteca comunale Filippo Maria de Sanctis - Cecosano - FR
☐	LC0035	LO131	Biblioteca civica Uberto Pozzoli - Lecco - LC

Licenza Attribuzione CC-BY 3.0

Condividi

15.40 27/09/13

The conversion

- ❑ To select a first set of records of modern material (“core elements”) and create an element for every allowed combination of tag, indicators, and subfield.
- ❑ Representation at finest level of granularity to be captured at the level of the UNIMARC subfield code and in the case of value vocabularies, the notation code and corresponding term, to avoid loss of information
- ❑ Intellectual value of UNIMARC is preserved by a finest-grained semantic representation



- Open Metadata Registry
 - <http://metadataregistry.org/>
 - FRBR/FRAD/FRSAD family
 - <http://iflastandards.info/ns/fr/>
 - ISBD
 - <http://iflastandards.info/ns/isbd/>
 - CIDOC-CRM
 - <http://cidoc-crm.org>
 - RDA
 - <http://rdvocab.info/>
- Linked Open Vocabularies
 - <http://lov.okfn.org/dataset/lov/>



Value vocabulary

- Dewey Decimal Classification

<http://dewey.info/>

- Nuovo Soggettario (BNCF)

<http://thes.bncf.firenze.sbn.it/>

- Resource Description Framework

Data format that supports simple, single metadata statements known as triples

- Each statement is in 3 parts
- Based on description logic
 - subject-predicate-object statements
- Also specifies relationships between things
 - thing-relationship-thing statements
 - Can be used for navigating between, or integrating, information from multiple source

- FRBR is a conceptual model built on a E-R methodology which is intrinsically applicable to representation in RDF, while ISBD is a data standard
- Design of the RDF representation of ISBD involves:
 - the treatment of aggregated statements in a defined number of elements within the areas;
 - the treatment of mandatory and optional elements and areas;
 - the order of areas and elements within an area;
 - the repeatability of areas and elements;
 - the treatment of punctuation and its double function.



FRBR and CIDOC-CRM

- FRBR/CIDOC CRM Harmonisation
- Expressing the IFLA FRBR model with concepts and tools provided by the CIDOC CRM
- Aligning the two object-oriented models
- Semantic interoperability between library and museum information

FRBR and CIDOC-CRM

- **FRBRoo**
- Formal ontology
- To capture and represent the underlying semantics of bibliographic information
- To facilitate the integration, mediation, and interchange of bibliographic and museum information
- Service Providers and Content Providers domains
- Well established and widely used standard
- Systems more stable
- Information more sharable

- W3C Semantic Web tools
 - <http://www.w3.org/2001/sw/wiki/Tools>
- W3C LinkingOpenData
 - <http://www.w3.org/wiki/SweoIG/TaskForces/CommunityProjects/LinkingOpenData>
- Open Metadata Registry (for vocabulary management)
 - <http://metadataregistry.org/>
- W3C SKOS
 - <http://www.w3.org/2004/02/skos/>

RDF Triples

- Metadata expressed as “atomic” statements
 - A simple, single, irreducible statement
 - The title of this book is “Cataloguing is fun!”
- Statements Constructed in 3 parts
 - “Triple”
- The title of this book is “Cataloguing is fun!”
 - Subject of the statement = *Subject*: This book
 - Nature of the statement = *Predicate*: has title
 - Value of the statement = *Object*: “Cataloguing is fun!”
- This book → has title → “Cataloguing is fun!”
 - *subject* → *predicate* → *object*



Triples from UNIMARC records

- Create or obtain URI for the Resource described
- Obtain URI for UNIMARC tag/subfield
 - Direct from tag/indicators/subfield encoding
- Obtain URI of value of subfield, or use a literal value
 - URI from vocabulary or UNIMARC Authority
- Publish triple

Related standards: Other

- The National Library of Sweden has developed a methodology for representing **MARC21 records in RDF** and implemented it for LIBRIS, the Swedish Union Catalogue
- The Vocabulary Mapping Framework (VMF) project [funded by UK Joint Information Systems Committee (JISC)]:
 - to develop a major expansion of the RDA/ONIX framework for resource categorization
 - to create a tool to support the **automated mapping of vocabularies from metadata standards** of use to the JISC community, which includes research, teaching, and learning environments
 - CIDOC CRM, FRAD, FRBR, MARC21 and RDA vocabularies [included] & ISBD and UNIMARC [represented]



UNIMARC to FRBRoo Mapping Example

001 IT\ICCU\TO0\1802857
005 20110418121539.2
010 \$a88-389-2536-4
010 \$a978-88-389-2536-8
020 \$aIT\$b2011-882
100 \$a20110304d2011 ||||0itac50 ba
101 \$aita
102 \$aIT
200 1 \$aRacconti per una sera a teatro\$fLuigi Pirandello\$ga cura di Guido Davico Bonino
210 \$aPalermo\$cSellerio\$d2011
215 \$a410 p.\$d17 cm
410 0\$1001IT\ICCU\CFI\0000165\$12001 \$a<<La >>memoria\$v844
676 \$a853.914\$cNARRATIVA ITALIANA, 1945-1999\$v21
700 1\$aPirandello\$b, Luigi\$f <1867-1936>\$3IT\ICCU\CFIV\000270\$4070
702 1\$aDavico Bonino\$b, Guido\$3IT\ICCU\CFIV\003294
790 1\$aBirandillu\$b, Luwgi\$3IT\ICCU\IEIV\076082\$zPirandello, Luigi <1867-1936>
790 1\$aPirantelo\$b, Louitzi\$3IT\ICCU\PUVV\323459\$zPirandello, Luigi <1867-1936>
790 1\$aPirandelo\$b, Luitzi\$3IT\ICCU\PUVV\323462\$zPirandello, Luigi <1867-1936>
801 3\$aIT\$bICCU\$c20130610
[...]





```
<frbr:F3_Manifestation rdf:about="http://www.sbn.it/IT\ICCU\TO0\1802857">
```

```
<crm:P102_has_title>
```

```
<crm:E35_Title>Racconti per una sera a teatro</crm:E35_Title>
```

```
</crm:P102_has_title>
```

```
<frbr:R9_comprises>
```

```
<frbr:F2_Expression rdf:about="http://www.sbn.it/expression/IT\ICCU\TO0\1802857">
```

```
<crm:P72_has_language>
```

```
<crm:E75_Language>ita</crm:E75_Language>
```

```
</crm:P72_has_language>
```

```
</frbr:F2_Expression>
```

```
</frbr:R9_comprises>
```

```
[...]
```

```
</frbr:F3_Manifestation>
```

Thank You!

Contact

Patrizia Martini

patrizia.martini@beniculturali.it

Achille Felicetti

achille.felicetti@gmail.com

www.athenaplus.eu

Twitter | LinkedIn (<http://linkd.in/19XwWOz>)

