

# The CoBiS Linked Open Data Project and Portal

Luisa Schiavone<sup>1,\*</sup>, Federico Morando<sup>2</sup>, and The CoBiS Communication Working Group<sup>3</sup>

<sup>1</sup>INAF Osservatorio astrofisico di Torino, Via Osservatorio 20, 10025 Pino Torinese, Italy

<sup>2</sup>Synapta Srl, Via S. Quintino 31, 10121 Torino, Italy

<sup>3</sup>[www.cobis.to.it](http://www.cobis.to.it)

**Abstract.** The CoBiS is a network formed by 65 libraries. The project is a pilot for Piedmont that is aiming to provide the Committee with an infrastructure for LOD publishing, thus creating a triplification pipeline designed to be easy to automate and replicate. This is being realized with open source technologies, such as the RML mapping language or the JARQL tool that uses Linked Data to describe the conversion of XML, JSON or tabular data into RDF. The first challenge consisted in making possible the dialog of heterogeneous data sources, coming from four different library software (Clavis, Erasmo, SB-NWeb and BIBLIOWin 5.0web) and different types of data (bibliographic, multimedia, and archival). The information contained in the catalogs is progressively interlinked with external data sources, such as Wikidata, VIAF, LoC and BNF authority files, Wikipedia and the Dizionario Biografico degli Italiani. Partners of the CoBiS LOD Project are: National Institute for Astrophysics (INAF), Turin Academy of Sciences, Olivetti Historical Archives Association, Alpine Club National Library, Deputazione Subalpina di Storia Patria, National Institute for Metrological Research (INRIM). The technical realization of the project is entrusted to Synapta, and it is partially sponsored by Piedmont Region.

## 1 Introduction

The CoBiS (“Coordinamento delle Biblioteche Speciali e Specialistiche di Torino” i.e. Coordination of Special and Specialized Libraries of Turin) is an informal network of 65 libraries, collaborating to provide continuing professional development and to offer a better service to their users.

CoBiS libraries are heterogeneous from many points of view: holdings, cataloguing softwares and OPACs. The **LOD project** started in 2015 as a training program, in collaboration with Prof. Vivarelli from the University of Turin. The program was divided in various topics: copyright, collaboration between libraries and Wikipedia, and Open Data.

## 2 Purpose

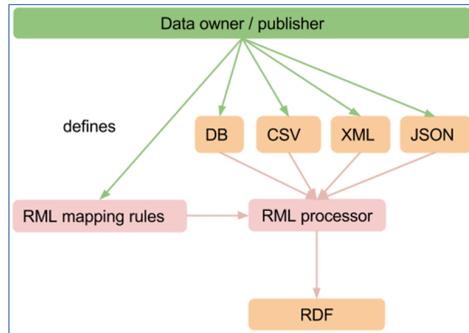
Turin Astrophysical Observatory Library and five other libraries from the CoBiS decided to participate in a **pilot project** with the purpose **to provide a unique access point to the collections of CoBiS libraries.**

---

\*e-mail: [luisa.schiavone@inaf.it](mailto:luisa.schiavone@inaf.it) ORCID: 0000-0003-2929-3727

CoBiS bibliographic data are not only becoming more interoperable among them, but the data are also opened in order to facilitate the collaboration with online communities and to connect them to the Linked Open Data Cloud.

During the project, **Synapta** made some improvements to the RML Processor to speed-up (by about two orders of magnitude) the triplification process.

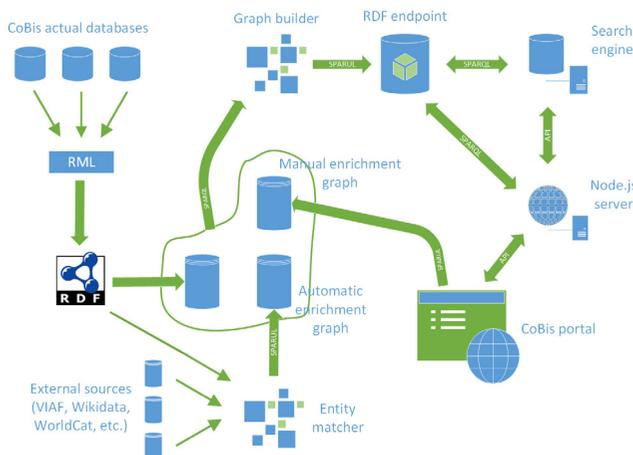


In order to further optimize the process, Synapta – in a joint effort with other open source developers – also realized and published JARQL (<http://jarql.linked.solutions/>), a new open source software tool for converting JSON data to RDF, using SPARQL 1.1 syntax, and using construct queries (see the project website or GitHub page for more details).

### 3 The Project

We used two main **ontologies**: Bibframe and Schema.org. We also used selected properties from RDFS, OWL, DCTerms, FAOF, and Culturalis (<http://culturalis.org/> – used to describe CoBiS libraries as cultural heritage institutions).

With respect to the **Linked Data stack**, this is an overall picture of the **project**



We started from heterogeneous sources, with the aim of combining them into RDF data. To reach this purpose, we used both **RML** (in particular in a first phase of the project) and our newly developed

**JARQL library.** Using a SPARQL query, we describe the transformation, obtaining an RDF file as result.

During the graph building, we also enriched data with external sources. We used both automatic algorithms and manual approaches.

For the latter, we exploited **OLAF (Open Linked Authority File), our crowd-sourcing interface for creating an authority file.** For instance, we want to match the correct Wikidata entity to a person named “Galilei”, who wrote the books mentioned on the right side. Calling on the Wikidata SPARQL API, we obtain different entities that you can choose to identify the exact matching. In this case Galileo Galilei.

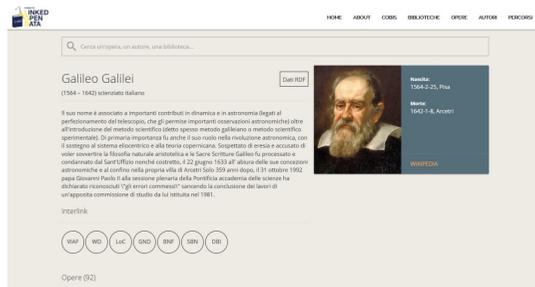


You may try OLAF live at <https://olaf.synapta.io/> selecting “CONNETTI” near “Autori dei libri del CoBiS” and register, in order to annotate the suggested triples with your identity. All the interlinks, discovered via OLAF and the other enriching techniques, allow the portal to access external resources: VIAF, Wikidata, the Italian National Catalogue SBN, Enciclopedia Treccani.

## 4 The Portal

Finally, as result, we have exposed data through a **SPARQL endpoint** that feeds the CoBiS portal <http://dati.cobis.to.it> (will be online by Fall 2017).

The CoBiS LOD Project portal is online with its full Linked Data stack, including a public SPARQL end-point (configured to support federated queries), a full dump of the RDF data, and Lod-View to dereference URIs, etc.



In the **author’s page - dynamically generated through SPARQL queries** – you have biographical information and a list of interlinked resources coming from Wikidata and other bibliographic

repositories (VIAF, Wikidata, LoC, Deutsche National Bibliothek GND, Bibliothèque Nationale de France BNF, Servizio Bibliotecario Nazionale SBN, Dizionario Biografico degli Italiani DBI).

On the right, an infobox with the author’s data and an image, both coming from Wikipedia, thanks to the leveraging of Linked Data. Clicking on the **RDF button**, all the triples of the resource can be directly fetched.

Alcuni scritti inediti di Galileo Galilei tratti dai manoscritti della...	Tomo 2 1714	Vol. 11 1900	Galileo Galilei... Sistema cosmico. In due dialogi 4. de dubio m...
Le opere di Galileo Galilei	3. Il terzo libro dell'opera burlesca del Berni del Galileo del Casa...	Vol. 8 1898	Vol. 9 1899
Vol. 2 1901	Tomo 10 1933	Supplemento Galileo Galilei	Tomo 7 1928
Le opere di Galileo Galilei: ristampa della edizione nazionale sotto...	Il ragionato di Galileo Galilei: [a cura di G. Chiarini]	1: Dell'anno 1567 sino alla fine del 1616	I dialoghi sui massimi sistemi tolemaico e Copernicano di Galileo Galilei...
Le opere di Galileo Galilei: edizione nazionale sotto gli auspici di Su...	Tomo 11 1934	Dialogo sopra i due massimi sistemi del mondo tolemaico e copernicano.	Vol. 8 1901
		Opere di Galileo Galilei	

At the bottom of the page, all the author’s **books inside the CoBiS database** are shown.

To explore information on such books, you can click one of the boxes or you can use the search bar and look for a title which is not listed.

Dialogo di Galileo Galilei Linceo matematico supremo dello studio di Padova, e Pisa... Dove ne i congressi di quattro giornate si discorre sopra i due massimi sistemi del mondo tolemaico, e copernicano; ...

Galileo Galilei | Ciccarelli, Lorenzo

Fronti, stampato in rosso e nero

Tipologia: singolo

Estensione: In questa seconda impressione accresciuta di una lettera della stessa, non più stampata, e di molti trattati di più autori i quali si reggono nel fine del libro...

Lingua: ITA

Luogo di pubblicazione: In Firenze in Napoli

Anno: 1710

Editore: In Firenze in Napoli 1710

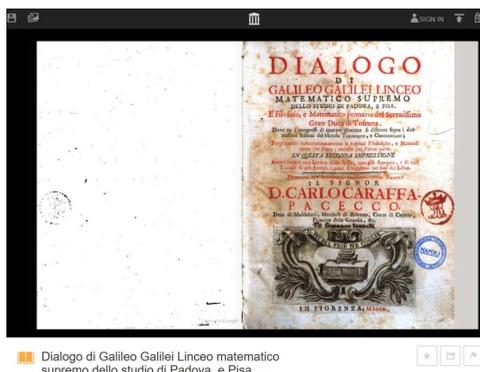
Dimensioni: 4"

Identificatore: 12.405.12.911 in B1 - 1 p.

OPAC

Biblioteche (1)

The image shows an **example search for the Dialogo**. On the left side of the page, you see bibliographic details with a collection of interlinked resources.



Exploiting the power of Linked Data, we are also able to read the **Internet Archive digital copy of the book**. A physical copy of the book is available in some CoBiS libraries. All details can be shown by clicking the **OPAC button**.

## **Acknowledgements**

1. Regione Piemonte for the financial support
2. Synapta for the technical realization
3. Nexa Center for Internet & Society at Politecnico di Torino for the collaboration
4. All the participating Institutes: National Institute for Astrophysics (INAF), Turin Academy of Sciences, Olivetti Historical Archives Association, Alpine Club National Library, Deputazione Subalpina di Storia Patria, National Institute for Metrological Research (INRIM).