Linked Data and cultural heritage data: an overview of the approaches from Europeana and The European Library

Nuno Freire
Data Services, The European Library
E-mail: nuno.freire@theeuropeanlibrary.org, nfreire@gmail.com

Europeana[1] provides access to digital resources from a wide range of cultural heritage institutions all across Europe. In order to support Europeana, a wide network of organizations collaborates in data integration activities. Within this network, The European Library[2] plays the role of library-domain aggregator, although its complete mission extends beyond cultural heritage into being a gateway to the collections and data of Europe's national and research libraries, operating on the principle of open data for re-use.

The Europeana Network addresses its data integration challenges by leveraging on Linked Data and the Semantic Web. Its approach to data integration is based in a single data model the Europeana Data Model, which embraces the Semantic Web principles to integrate, in an open environment, the various data models and ontologies used in cultural heritage data.

The paradigm of Linked Data, brings many new challenges to libraries. The generic nature of data representation used in Linked Data, while it seamlessly allows any community to manipulate the data, also brings many open paths to its implementation, with no clear optimal choice for libraries. The European Library leverages on its operational infrastructure to make library data available as Linked Open Data. It maintains The European Library Open Dataset, which is derived from the data aggregated from member libraries, and made available under the Creative Commons CC0 1.0 Universal license, in order to promote and facilitate the reuse of the data by libraries and other communities.

Extensive linking is performed in the preparation of The European Library Open Dataset. It relies on Information Extraction and Data Mining techniques to establish links to external linked open datasets, covering the most prominent entities types present in library data and: persons, corporate bodies, places, concepts, and intellectual works and manifestations.

The European Library has also an innovative linked data approach for intellectual property rights clearance processes, used for supporting mass digitization projects. This approach is applied in the within the European ARROW rights infrastructure[3].

[1] http://www.europeana.eu

[2] http://www.theeuropeanlibrary.org

[3] http://www.arrow-net.eu/