ALIADA, an Open Source Tool for Automatic Publication of Linked Data from Libraries and Museums

Cristina Gareta

ALIADA Project Consortium, Vitoria-Gasteiz, Spain
cgareta@scanbit.net
ALIADA Project Consortium, Rome, Italy
ALIADA Project Consortium, Budapest, Hungary

Abstract. ALIADA, Spanish word that means ‘ally’, is intended to be a tool to help librarians and curators from cultural heritage institutions to automatically publish their high quality data in the Linked Data Cloud. Traditionally, data from libraries and museums have been stored as ‘silos’ of information because their metadata are codified using their own schemes and formats, not accessible by machines and applications more general public-focused. In addition, these information professionals create rich data from their collections, but they are not expert enough to face the coming technologies required to take advantage of the opportunities that the information and open knowledge era provides. To overcome these limitations, ALIADA EC-funded Project has developed an open source tool compliant with libraries and museums standards that automatically converts library and museum metadata into structured data ready to be published in the Linked Data Cloud, according to the Linked Data paradigm. Thus, heritage and cultural data are also open and available to be queried and reused by machines, innovative applications, search engines and other cultural and research institutions to generate more open knowledge.

Keywords: Libraries · Museums · Linked data · Semantic web · Open data

1 Introduction

Memory institutions or GLAM (Galleries, Libraries, Archives and Museums) have to face the challenge of opening their data in the era of the open knowledge society. They preserve cultural heritage and provide universal access to information about it to people, communities and organizations. These institutions generate high quality local information particularly interesting to the web community because often it is not accessible by web search engines and, therefore, it is not open to the world. This is because institutions such as museums and libraries have their own data codification standards. As services of public bodies, they should publish their datasets as “linked data” to enrich the web and to be reused to create new knowledge, to help other cultural institutions to be more efficient or to contribute to innovative applications.

In this context, ALIADA Project is an European Commission-funded project that aims to develop an open source tool to be the “ally” (that is the translation in Spanish)
for non-expert users, such as librarians and museum curators, in the publication of their datasets in the Linked Open Data Cloud. This tool is intended to be multilingual and suitable to integrate in existing library and museum systems to overcome the limitations of those systems. ALIADA tool supports the current standards for cataloguing bibliographic and authority records in libraries (MARC, Dublin Core) and collections and objects in museums (LIDO). Also the standards for the Semantic web technologies are used in this tool: RDF, URI, OWL, HTTP or SPARQL.

The project was promoted by IT companies experts in library software (Scanbit, in Spain, and @Cult, in Italy), art museums (Artium, in Spain, and Museum of Fine Arts Budapest) and a research agency expert in Semantic web technologies (Tecnalia). ALIADA was born as a two-year project and it is expected to be finished in October 2015. After the first year of project, ALIADA Consortium has launched the first prototype which is available for downloading on GitHub and on ALIADA website.

2 Project Goal and Objectives

Linked data (often capitalized as Linked Data) is a method for exposing, sharing, and connecting pieces of data, information, and knowledge on the Semantic Web using URIs and RDF. To help libraries and museums to publish their data as “linked open data” (linked data that is open content), ALIADA project aims to provide an open source tool compliant with the library and museums' data management systems and with the Linked Data Cloud rules and standards. According to that goal, the following are the main objectives that guided the development of this innovative tool:

- Automatic publication in the Linked Open Data Cloud (LODC) of data generated by libraries and museums. Input data are descriptions of cultural and heritage collections stored and maintained by public institutions. Metadata schemes used for cataloguing those documents and objects are specific for libraries and museums (MARC, LIDO)
- Development of a Java-based web application to integrate with library and museum management systems, as a plugin or as an external tool supporting their metadata schema and file formats
- ALIADA will be a non-expert user oriented application. Librarians and curators are experts in cultural information and heritage management, but they don't know about the Linked Open Data technology. Therefore, usability will be a key aspect.
- Open data requires open source code. It's expected the creation of a community of developers and users around the ALIADA open source tool. Target users are libraries, museums, research agencies, IT SMEs and developers.
- As European project, ALIADA will be multilingual. Initially, it will support English, Spanish, Italian and Hungarian languages.

1 http://aliada-project.eu
2 http://linkeddata.org/