ASSETS: Advanced Service Search and Enhancing Technological Solutions for the European Digital Library

Nicola Aloia, Cesare Concordia, and Carlo Meghini

Institute for Information Science and Technologies of the Italian National Research Council (ISTI-CNR), Pisa
{nicola.aloia,cesare.concordia,carlo.meghini}@isti.cnr.it

Abstract. ASSETS is a 2 year project co-funded by the CIP Policy Support Programme. The main goal of the project is to improve the usability of Europeana (the European Digital Library) by developing, implementing and deploying largescale services focusing on search, browsing and user interfaces. ASSETS strives also to make more digital items available on Europeana by involving content providers across different cultural environments.

1 Introduction and Motivation

Heritage, history and cultural diversity represent all together an enormous yet dispersed strength which unanimously contribute to create the Europe uniqueness. Europes libraries, museums, archives and audiovisual archives possess huge and rich collections, which, once made accessible by the Internet, allow people to consult and reuse them for several different purposes, like work,leisure and study, and for a wide range of information product and services (e.g. tourism, education). European Commission has been acknowledging the importance to preserve, maintain, sustain and make European cultural heritage accessible for some years. In particular the Digital Libraries initiative, which is flagship project of the EU Commission i2010 strategy for the Information Society, is strongly supporting the setting up and the development of Europeanathe European Digital Library, which aims at making accessible to a wide public books, films, newspapers, photographs, drawings and archival documents from European cultural institutions, through a unique web portal.

The ASSETS project aims at developing digital library services focusing on search, browsing and interfaces in order to improve the usability of Europeana. ASSETS will implement and deploy largescale services for:

- Searching multimedia objects based on metadata and on content similarity;
- Browsing multimedia objects for rapid navigation through semantic crosslinks;
- Building Interfaces designed for interacting with multimedia objects;
- Planning Longterm Access to digital information;
- Ingestion of metadata requiring normalization, cleaning, knowledge extraction and mapping to common structure.

M. Agosti et al. (Eds.): IRCDL 2011, CCIS 249, pp. 104–106, 2011.

The main expected outcomes from ASSETS project can be summarized as follows:

- a set of valueadded services to directly enhance the content accessibility and improve the usability of Europeana
- a set of operational services adding new multimedia digital content to Europeana,
- a set of recommendations, guidelines and best practices concerning the technology adaptation, and service rolling up, to contribute to the promotion of the standards and specifications addressed in Europeana, and pave the way for an increased interoperability.

2 Scientific Challenges

From a scientific point of view, participating to ASSETS gives to researchers the opportunity to have access to very high quality data, such as those made available to the project by the involved stakeholders; and to have direct contact with the users and the stakeholders, thus gaining requirements for new challenges.

3 Key Technologies

Considering the "three-tier architecture" paradigm typically adopted to describe information management systems, the ASSETS system can be described as follows:

Presentation tier: The ASSETS system will provide RESTful APIs enabling external applications to use its functionalities, a set of GUIs for managing specific functionalities (e.g. data ingestion and access) and a web GUI (ASSETS portal) for browsing services and functionalities.

Business logic: A set of autonomous software components implementing the functionalities described in the previous paragraph interacting each others using an ad hoc integration component.

Data tier: An heterogeneous set of repository managers to manage different kind of information objects: multimedia objects, structured data, etc.

The ASSETS system is a loosely coupled system: each software component implementing the business logic has no (or little) knowledge of other components and also is not supposed to know which tools are actually used to store or manage data.

The ASSETS system rely on the top of the technology outcomes of the PHAROS, MESH, BRICKS, CASPAR, PLANETS and also other European related projects such as eConnect.

4 Contribution by Italian Research Community

CNR-ISTI has a large participation in the ASSETS project.

The ISTI team is responsible of several fundamental activities and is involved in all R&D activities. The main ISTI activities are:

- **Ingestion of metadata and machine learning:** The ISTI team will deliver a suite of services implementing sophisticated functionalities for metadata cleaning, knowledge extraction and data classification. The goal of these services is to improve the quality of the content stored in Europeana.
- **Indexing and retrieval of data:** A new search service for Europeana will be provided. The service will use advanced indexing and query processing techniques in order to return the most significant result set and will also provide a similarity-search functionality for multimedia objects.
- **User Generated Content and personalization:** The ISTI team, in cooperation with the University of Paris Sud, will deliver a service managing the ingestion of user generated information objects in Europeana. This service will enable users to create and manage their own objects also in a cooperative way (notifying actions to other users etc) and to submit the object to Europeana, if accepted those objects will be part of the Europeana information space.

Another important activity of ISTI is the collaboration to the integration of ASSETS development platforms with the Europeana Labs.