

# EuropeanaConnect

## Enhancing user access to European digital heritage

Max Kaiser  
Austrian National Library  
Email: [max.kaiser@onb.ac.at](mailto:max.kaiser@onb.ac.at)

Jeanna Nikolov-Ramírez Gaviria  
Austrian National Library  
Email: [jeanna.nikolov@onb.ac.at](mailto:jeanna.nikolov@onb.ac.at)

Veronika Prändl-Zika  
Austrian National Library  
Email: [veronika-praendl-zika@onb.ac.at](mailto:veronika-praendl-zika@onb.ac.at)

### Abstract

This paper provides an outline of the main goals of the project EuropeanaConnect, one of the two core technical projects in the Europeana project cluster which will develop the current prototype of Europeana into a fully operational service until 2011. Europeana is Europe's digital library, archive and museum portal, launched in November 2008.

**Keywords:** Europeana, semantic web, multilingual access, mobile access, value-added services, licensing framework, audio content

### 1 Introduction

The prototype of Europeana (<http://www.europeana.eu>), Europe's digital library, archive and museum portal, was launched in November 2008. The European Commission is funding the development of Europeana as part of its i2010 Digital Libraries initiative to create an integrated access channel to Europe's distributed digitised cultural heritage resources.

The launch of the Europeana prototype created such a large-scale public interest that the site had to be taken down for a few weeks for a substantial hardware and load balancing reconfiguration. Back online in a test mode since December 2008 and with full functionality since April 2009, the Europeana prototype currently gives access to about five million digitised objects of cultural value from over 1.000 European libraries, archives, museums and audio-visual collections.

Within the next two years Europeana will be developed into a fully operational service. By mid 2010 Europeana is expected to provide access to 10 million items; the target for 2012 is 25 million items. The European Commission is funding a number of projects in the eContentplus and ICT PSP programmes. Within the next years, these projects will develop the technology and services for Europeana and digitise and aggregate content for the portal.

Two core projects will be responsible for the implementation of the technical infrastructure of Europeana: The project *Europeana v1.0* (<http://version1.europeana.eu/>) is led by the European Digital Library Foundation. This 30-month project which started in February 2009 will solve key operation issues related to the implementation of Europeana like building the

back end services needed to manage the delivery and access of Europeana content and managing the channels that will enable other environments to use the content made interoperable by Europeana via web services and APIs. This project will also develop and manage partnerships to ensure a rich content flow from national and domain aggregators. The second core project is EuropeanaConnect (<http://www.europeanaconnect.eu>) and will be described in the remainder of this paper.

## 2 EuropeanaConnect

EuropeanaConnect is closely associated with the project *Europeana v. 1.0* and is a 30-month Best Practice Network co-funded by the European Commission in the *eContentplus* programme. Co-ordinated by the Austrian National Library, EuropeanaConnect will deliver core components and value-added services for Europeana. It will also act as a content aggregator for audio content and will significantly enlarge the music content available in Europeana. The project runs from May 2009 to October 2011 and brings together 30 partners from 14 European countries, including universities, research institutes, libraries, audio archives and a partner from the publishing industry.

### 2.1 Europeana Semantic Layer

One of the main goals of EuropeanaConnect is to build a layer of semantic data which will be the basis for all semantic processing in Europeana. Semantic processing would, for example, enable a relationship between “Mona Lisa”, “Lisa del Giocondo”, “La Joconde” and “Leonardo da Vinci” to be recognised during the search process. A user entering just one of these terms would see also results for the linked concepts, e.g. would automatically find other paintings created by Leonardo da Vinci. EuropeanaConnect will provide the technologies to semantically enrich huge amounts of digital content in Europeana. This will enable semantically-based content discovery and make Europeana content more accessible, reusable and exploitable.

EuropeanaConnect will create a network of semantic resources starting from controlled vocabularies and other similar resources which will be provided by content owners. These will be made ready for semantic processing mostly by using SKOS (Simple Knowledge Organization System, a W3C standard) which in turn will be used as the primary level of user interaction with Europeana. Unlike in traditional library catalogue-driven models of user interaction, users of Europeana will be able to explore the data space by using the semantic nodes (i. e. concepts organised and available as web resources) as primary elements for searching and (faceted) browsing.

### 2.2 Multilingual Access

Multilingual and multicultural aspects are at the heart of making Europeana’s digital contents effective and exploitable for users across all European countries. Metadata and digital content should be searchable and presentable independent of the language of the searcher or the object descriptions. In other words, users should be able to enter search terms in their own or preferred language, which are translated on the fly during the search process, and have hits returned in a range of selected languages. In order to enhance the multilingual access capabilities of Europeana, EuropeanaConnect will develop a multilingual infrastructure and a set of translation tools that will process queries and object data and produce a suitable multilingual representation for the user:

- a unique repository of language and translation resources for multilingual processing of objects within Europeana, which will possibly be one of the most varied of its kind in Europe;
- a suite of tools for multilingual mapping of controlled vocabularies in at least five European languages and a strategy for incorporating more languages;
- a suite of integrated translation modules for querying and browsing processes which will plug into the general search and retrieval infrastructure of Europeana.

Not every European language is equally well developed in terms of multilingual resources for translation. EuropeanaConnect will start with a core set of languages (English, French, German, Italian, Spanish) for which expanded multilingual translation capabilities will be implemented by the end of the project. During the course of the project, the core set will be expanded, based on the development experiences, with a set of additional languages integrated with less linguistic treatment based on available resources and content.

### 2.3 Tools and methodologies for user-driven development

EuropeanaConnect will provide a better understanding of what current and future web-users really need and expect of the multi-lingual and multi-channel service which Europeana must become. The project will provide the methodology to monitor and evaluate the user interactions with Europeana. Deep Log Analysis will allow for the evaluation of significant information-seeking behaviour of Europeana users. This method will be used by Europeana as a routine for understanding how the Europeana portal is used and what its users really expect from it. The usage-monitoring methodologies and the results of Deep Log Analysis will inform the implementation iterations of the Europeana portal (*Europeana v1.0*) and will thus ensure that the development of Europeana is truly user-driven.

Europeana sees all groups of people as potential users. It will not be sufficient to have information on the actual user behaviour; we also need to move from the concept of the user as “anonymous” to an understanding of who is using what. To support this and to generally ensure user-oriented service development and marketing of Europeana, EuropeanaConnect will define “personas”, i.e. categories with profiles. We need to understand the link between the results of the deep log analysis and these personas and we need to understand how Europeana can add value to a selected number of these personas. This work will be accompanied by actual user studies either through questionnaires or via different forms of usability tests. From a preliminary screening we aim for a rough number of 500 people to be involved throughout the lifetime of the project.

### 2.4 New Access Channels to Europeana

#### 2.4.1 Spatio-temporal user interface

EuropeanaConnect will develop a spatio-temporal access service that will support access to Europeana via a number of space- and time-related selection mechanisms as part of the query- and result-presentation process. It will implement a user-friendly access mechanism, which exploits the Semantic Layer of Europeana. The objective is to create an interface that makes use of both time-based and geographical metadata of the Semantic Layer, providing a new visual access to the digital collections in Europeana. The project will build an interactive and generic map combined with a timeline, acting as a “histogram” for tagged events, places and characters in Europeana. The interactive map will show the visual representation of a search-result set in Europeana, e.g. concerning the search term “migration”. The histogram on

the timeline below the map will indicate how many items are registered in Europeana within a specific timeframe.

#### 2.4.2 Mobile Access Channel

EuropeanaConnect will address the user demand to access information through mobile devices. The project will develop both a middleware layer, which will allow different mobile devices to access Europeana services created and a rich mobile client for real end user evaluation. The middleware will allow for access to Europeana from different mobile devices, channels and applications. It will be located between mobile devices on the user side and the Europeana semantic layer and database on the server side. It is intended to overcome the heterogeneity and limitations of mobile devices by selecting and adapting the content for queries to make them displayable on the devices. Therefore, queries from mobile devices are processed by the middleware, which uses the Europeana semantic layer to retrieve the desired data. The query results are collections of media which have to be adapted for mobile usage and a presentation of the results based on the metadata will be prepared. Finally, this presentation will be transferred to the mobile device. Advanced queries and interfaces require more powerful mobile devices, such as Smartphones. Therefore EuropeanaConnect will also implement a rich mobile client application which runs on Windows Mobile-based devices.

### 2.5 Value-Added Services for Europeana

EuropeanaConnect will integrate three value-added services with the Europeana portal that will significantly enhance the functionality and the usability of Europeana:

#### 2.5.1 Multimedia Annotation Service

This service will allow all Europeana users to make their own contributions to Europeana in the form of tagging, comments, discussions, and linking. Annotations are an important means by which users can contribute knowledge and exchange information about content within a group of experts. Cultural institutions and museums that have already exposed their content on the Web are discovering that annotations can change the traditional role of visitors from passive consumers to actively contributing and collaborating users. Annotations allow users to exploit and add value to existing content, while enhancing accessibility through the user-provided searchable metadata and links to other media. Annotations further enable communication and collaboration with other users interested in the same content, which will support the establishment of social networks built around Europeana. The Annotation Service will allow the tagging, linking, and annotating of Europeana media resources of various formats (HTML, image, audio, and video).

#### 2.5.2 GIS Service Suite

The GIS Service will allow users to query and display Europeana content based on spatial information, and to discover new relationships between content items, based on location. As only very few objects in Europeana already contain explicit geographical metadata, this service will also provide metadata enrichment for the spatio-temporal visual user interface described above. EuropeanaConnect will integrate a Geoparser and Gazetteer developed within the context of the DIGMAP and TELplus projects with the Europeana portal. By processing Europeana metadata with the Geoparser it will be possible to detect geographic terms and time periods. This will allow to improve the Europeana metadata. Pre-processed data will be stored in the semantic layer and will later be used to populate the spatio-temporal user interface. Finally, we can apply the same processing in real time for the users' queries. When the Geoparser finds a relevant geographic or time reference (for example, "churches in

the Scotland”, “the battle of Lepanto”, etc.) in a query, it can trigger specialised services that can take that in consideration.

### 2.5.3 eBooks-On Demand

Although Europeana provides access to some millions of digitised items from European cultural heritage institutions, these figures represent only a very small fraction of the analogue collections available in these institutions. In the case of books we know that the portion of already digitised material is around one percent of all books, in the case of archival material it will be even lower. This situation has been the starting point for the eBooks on Demand (EOD) Service that will be integrated into Europeana: The EOD Network (<http://books2ebooks.eu/>) is a growing consortium of 18 libraries from 10 European countries. They all provide the EOD service which allows users to order on-demand digitisation of public-domain books of the participating libraries and to receive the requested items electronically in the PDF format. EuropeanaConnect will adapt the EOD service allowing for an automated and generic transfer of the eBooks created within EOD to Europeana according to the specification of the Europeana Semantic Elements.

## 2.6 Infrastructure components for Europeana

EuropeanaConnect will implement and deploy key infrastructure components that will enable Europeana to manage the harvesting of thousands of digital heritage content sources across Europe, to promote metadata and content interoperability across content providers as well as interoperability with independent value-added services, and to provide persistent and uniform identification of digital resources:

- The Europeana OAI-PMH Management Infrastructure will be a prerequisite for enabling Europeana to meet the demands of the expected leap in data harvesting scale, managing the harvesting of huge volumes of content from thousands of content providers in Europe.
- The Europeana Metadata Registry will be a precondition for metadata interoperability in Europeana, as it will manage the range of metadata schemata and terms used by the various European cultural heritage institutions
- The Europeana Service Registry will enable Europeana to integrate practically any external service offering an HTTP interface into the portal. It will allow users to search and select value-added services and will provide the functionalities for the Europeana portal to invoke appropriate services depending on the context.
- The Europeana Resolution Discovery Service for Persistent Identifiers will allow for uniform and persistent identification of resources between Europeana and service and content providers. It will provide a crucial component for the portal’s interoperability that supports the integration of a critical mass of European digital heritage content for the end user.

## 2.7 Europeana Licensing Framework

Even though content of Europeana is likely to be in the public domain or with the rights cleared, it will be essential to provide the user with an accurate description of the rights attached to the content. Appropriate licensing policies for objects and metadata as well as suitable implementation methods are thus needed in order to create trust for all parties. EuropeanaConnect will create the Europeana Licensing Framework and will build the necessary tools for Europeana to establish and declare the rights status associated with Europeana content. In close collaboration with the Europeana Office and the *Europeana v1.0*

project, EuropeanaConnect will establish a core set of interoperable licenses that cover the rights information for objects in Europeana. The proposed Europeana Licensing Framework will be validated through a broad stakeholder consultation process.

## 2.8 Aggregation of Audio Content

Apart from building critical components for Europeana, EuropeanaConnect will also act as content aggregator: One of the shortcomings of the current Europeana prototype is its lack of a critical mass of non-textual content. EuropeanaConnect will add the music dimension to Europeana. As well as harvesting massive amounts of audio, the project will create, implement and evaluate access to right-free or wholly-owned music for all users of Europeana. EuropeanaConnect will integrate the framework created by the DISMARC *eContentplus* project (<http://www.dismarc.org/>) in Europeana and extend the technical, editorial and legal solutions of this platform. The project will provide the necessary infrastructure for harvesting, analysing and storing of audio metadata as well as music audio provision to Europeana. There will be a continuous approach to and involvement of new audio archives for the duration of the entire project, using the established infrastructure and harvesting process. It is anticipated that over the period of the project approximately 150 audio archives will participate as content providers, leading to an estimated 200.000 music tracks to be provided to Europeana by month 30.

## 3 Conclusion

The launch of Europeana has yielded a high public interest. Within the next two years the current prototype has to be developed into a fully functional and scalable service that meets the high expectations of the user community. The project EuropeanaConnect which has been outlined in this paper will contribute to this goal on several levels. It will deliver key technical components which are essential for the realisation of Europeana as an interoperable and multilingual service. It will establish methodologies for analysis of user-interaction that will inform the implementation the Europeana portal is driven by user needs. Moreover it will deliver the Licensing Framework for Europeana and aggregate a critical mass of audio content.

## Acknowledgement

EuropeanaConnect is co-funded by the European Commission in the *eContentplus* programme.