

**2012 digitization awareness seminar**  
**Lagos, 25 April 2012**  
**Digital Library: Innovative Approaches to the Creation and Delivery of Digital Content**

---

**Overview of digitization trends in Europe**

Julie Verleyen, [verleyenj@yahoo.fr](mailto:verleyenj@yahoo.fr)

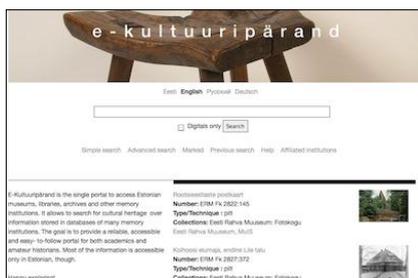
**Digitization landscape in Europe**

Since the beginning of the web and the automation of processes European institutions such as libraries, research centers, museums, universities, archives, etc... have been running projects to digitize the content they hold.

Plenty of digital collections of all types are now available through intranet and internet services. While the creation of new sets of digital resources goes on (old manuscripts, local heritage collections, thematic collections...), scaling up and improving existing digital collections is part of the agenda of most countries throughout Europe.

In the cultural sector, many countries realize their national portals: *Hispana* in Spain [<http://hispana.mcu.es>], *e-Kultuur* in Estonia [<http://e-kultuur.ee/?locale=en>] , *CulturalItalia* in Italy [<http://www.culturalitalia.it/>], BAM portal in Germany [<http://www.bam-portal.de>], CiMEC [[http://cimec.ro/e\\_default.htm](http://cimec.ro/e_default.htm)] in Romania, *Moteur Collection* in France [[http://www.culture.fr/fr/sections/collections/moteur\\_collections](http://www.culture.fr/fr/sections/collections/moteur_collections)], etc...

Beyond the national boundaries, the different domains collaborate to create common platforms, at European level: *The European Library* [<http://theeuropeanlibrary.eu>] for national libraries, *European Film Gateway* [<http://www.europeanfilmgateway.eu/>] for the audio-visual archives, *Archives Portal Europe* [<http://www.archivesportaleurope.eu>] for archives, ATHENA [<http://www.athenaeurope.org/>] for the museums.



^ National portal for Estonian Culture    ^ Portal of European Archives

^ European Film Gateway

New initiatives include, for example, *Europeana Regia* [<http://www.europeanaregia.eu>] which aims at digitizing circa 880 old manuscripts from the Middle Age and Renaissance periods. Five European libraries<sup>1</sup> work together to constitute a unique virtual library of documents that illustrate

the royal history in Europe of the Carolingian dynasty, the Charles V and family dynasty and the Aragonese kings of Naples dynasty.

We can also cite some thematic project, such as the *Newton Project* [<http://www.newtonproject.sussex.ac.uk>] which brings the works of Isaac Newton in a digital form, or *Musical Instruments Museum Online (MIMO)* [<http://www.mimo-db.eu/>], a pan-European collection of images and descriptions of musical instruments held in museums.

### ***Trends***

In October 2011, LIBER<sup>2</sup> and EBLIDA<sup>3</sup>, two major libraries associations in Europe, held for the 3<sup>rd</sup> time a workshop on the digitization of material in Europe. From this event<sup>4</sup>, it transpired that building the capacity, improving workflows and user experience, serving the digital content in innovative ways are the objectives of most digitization programs.

In overall, the emphasis is put on the following aspects:

- Collaborate and Innovate
- Enhance and improve
- Share and re-use

The development of networks and associations from all sectors – academic, cultural, scientific... – is also an important aspect of the digitization movement. These networks allow to connect digital repositories - in a centralized or distributed way – in order to grow the amount of digital items available on the web at once, for all web users and beyond boundaries.

Also, these networks and associations put together groups of experts and professionals: through collaboration platforms, they share competences and experiences. This has a positive impact, for example on the quality of new and existing collections, benefiting from the emergence and application of standards, good practice methods and tools.

### ***The context in Europe: political directive, funding structure***

The political role and funding schemes in Europe are particularly important for digital activities. For example, the European Commission (EC) makes recommendations on digitization and digital preservation. In the latest recommendation<sup>5</sup> that was adopted in October 2011, it invites the member states' partners to set clear targets and budgets for digitization, as well as overviews of materials that are already digitized and where they are accessible, at European level. It encourages the optimal use of digitization capacity, the reuse of existing supporting structures (e.g. competence centres) and the

---

1 The Bavarian State Library, the French National Library, the Belgian Royal Library, the historical library of València University in Spain and the German Herzog August library

2 LIBER [<http://www.libereurope.eu>] is the *Ligue des Bibliothèques Européennes de Recherche – Association of European Research Libraries*

3 EBLIDA [<http://www.eblida.org/>] is the *European Bureau of Library, Information and Documentation Associations*. It is an independent umbrella association of library, information, documentation and archive associations and institutions in Europe

4 [<http://indicotest.eblida.org/conferenceDisplay.py?ovw=True&confId=22>] The Workshop intends to take the idea of digitization of library material further by discussing what potential libraries have for developing Europe's digital agenda through four sessions: User Experience, Business Models, Reuse of Metadata and Workflow.

5

[http://ec.europa.eu/information\\_society/activities/digital\\_libraries/doc/recommendation/recom28nov\\_all\\_versions/staff\\_working\\_paper.pdf](http://ec.europa.eu/information_society/activities/digital_libraries/doc/recommendation/recom28nov_all_versions/staff_working_paper.pdf)

aggregation with digitization efforts. The recommendation also encourages public-private partnerships, under conditions such as the preservation of public materials in the public domain, the transparency of processes or the period of exclusive use that shouldn't exceed 7 years.

### ***Collaborate and Innovate***

Under thematic frameworks, a number of projects are funded by the Commission to realize the governmental objectives of the country members, at European level.

As part of the EC's Information Society segment, the Digital Libraries Initiative, in combination with the ICT-Policy Support Programme, initiate projects to digitize, preserve and give access to the European heritage.

In general, these projects are built upon existing local/regional/national initiatives, and they associate both the public and private sectors. Representative outcomes include:

1. the Open Planets Foundation<sup>6</sup>: Created in 2010, it is the community hub for digital preservation in Europe. The website features the contribution and expertise from practitioners in the field, with blog posts and forums. Based on standards and latest technologies, it promotes a software suite to implement an effective digital preservation infrastructure.
2. the Impact Centre of Competence<sup>7</sup>: Launched in October 2011, the centre provides a solid framework to support digitization through practical resources. The IMPACT's website hosts an organized knowledge bank – with guidelines, videos of presentations and tutorials, learning objects – as well as offline and offline tools. All phases and processes of a digitization workflow are covered. Upstream of this workflow, for example, the *Digitization Cost Estimator* is an excel-based tool to evaluate the costs involved for a project, taking into account the number of volumes and pages to be scanned, the estimated storage, the staff etc... Another example, down the workflow, is the *OCR Post-correction and Enrichment* toolkit. This toolkit integrates a tool, originally developed by the Stanford University, that recognizes named entities – words representing persons and places names - in a scanned text. Once extracted, these valuable words can be used for effective indexing and search purposes.
3. The Europeana Foundation and portal<sup>8</sup>: Since 2008, the Europeana web service is a single point of access for the digital cultural heritage of Europe. Today, users can search and browse through 20+ millions of objects from all countries, through a multilingual interface. Europeana is also a foundation that federates the efforts across domains to bring all European digitized content together, being an image, a text, a video or an audio recording. Here is an example of aggregation scenario in Europeana, based on the metadata harvesting mechanism OAI-PMH (Open Archive Initiative-Protocol for Metadata Harvesting): The Royal Museum of Arts and History in Belgium hosts a collection of musical instruments that can be viewed online, on the museum's website. This institution contributed to the previously mentioned MIMO project, which aggregated all digital collections of music instruments across

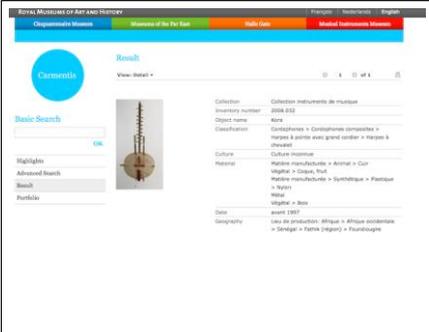
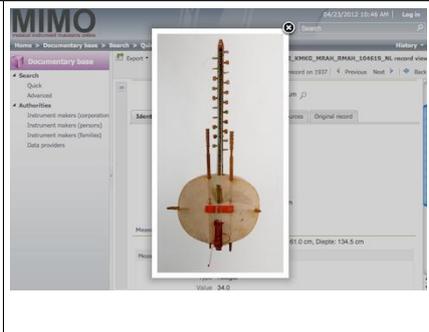
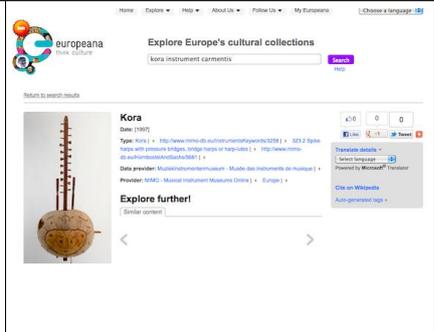
---

6 The Open Planets Foundation (OPF) [<http://www.openplanetsfoundation.org>] has been established to provide practical solutions and expertise in digital preservation, building on the €15 million investment made by the European Union and Planets consortium.

7 IMPACT Centre of Competence for Digitization [<http://www.digitisation.eu/>]

8 Europeana portal [<http://europeana.eu>] and Europeana Foundation [<http://pro.europeana.eu/web/guest/>]

Europe, into a thematic platform. In its turn, MIMO contributed the content of its repository (just the metadata) to Europeana which ingest it and offer it through its own website.

		
<p>1. Object at hosting institution's website: <a href="http://carmentis.kmkg-mrah.be/eMuseumPlus?service=ExternalInterface&amp;module=collection&amp;objectId=104619&amp;viewType=detailView">http://carmentis.kmkg-mrah.be/eMuseumPlus?service=ExternalInterface&amp;module=collection&amp;objectId=104619&amp;viewType=detailView</a></p>	<p>2. Object at thematic aggregator's website: <a href="http://www.mimo-db.eu/mimo/infodoc/Ged/View.aspx?eid=OAI_KMKG_MRAH_RMAH_104619_NL&amp;searchId=35485f20-a655-41aa-99e5-f478766a0339">http://www.mimo-db.eu/mimo/infodoc/Ged/View.aspx?eid=OAI_KMKG_MRAH_RMAH_104619_NL&amp;searchId=35485f20-a655-41aa-99e5-f478766a0339</a></p>	<p>3. Object at cultural aggregator's website: <a href="http://europeana.eu/portal/record/09102/40F1932A75BB33B8B5AA4C998553EC167489144E.html?query=kora+instrument+carmentis">http://europeana.eu/portal/record/09102/40F1932A75BB33B8B5AA4C998553EC167489144E.html?query=kora+instrument+carmentis</a></p>

It is relevant to outline partnerships with commercial companies, as a choice made by several large European public organizations for their digitization plans.

Libraries of the cities of Munich in Germany, Oxford in United Kingdom and Lyon in France made agreements with Google Books, the online books program of Google, to enrich their virtual collections with more digital volumes. The National Library of Austria, is currently realizing its Austrian Online Books project<sup>9</sup> with Google. The objective is to digitize 600.000 historical volumes representing the 16<sup>th</sup> - 19<sup>th</sup> century period. A specific workflow has been put in place to manage the physical books on one side, with specific preparation and manipulation conditions, and the data resulting of the digitization operations on the other side. While the current project allows to realize mass digitization in a relatively short period of time, the further enhancement of the digitized materials will be the object of another project. With the help of a software, the Austrian Book Online Download & Control, an automatic correction of systematic errors is applied. The refined correction of individual problems will be addressed at a later stage.

### ***Enhance and improve***

The increasing of digitized content also implies to build capacity around it and getting the participation of the public is an approach that we can observe in Europe. Crowd-sourcing is a strategy that is more and more taken to realise activities that require lots of human resources and /or that computers cannot yet perform.

For example, it is used to improve the quality of access to digitized content. The National Library of Finland has launched an original project to engage any web user to work on the quality of its old Newspapers collection, after OCRing. Digitalkoot<sup>10</sup> proposes to anyone who wants to participate to play games in order to complete and fix OCRed words. Players read words that the machine couldn't read properly and type them in. In that way, an imprecisely OCRed text gets corrected and

<sup>9</sup> <http://www.onb.ac.at/ev/austrianbooksonline.htm>

<sup>10</sup> <http://www.digitalkoot.fi/en/splash>

can be fully indexed, with accuracy.

The contribution of the public to create digital content itself is also an interesting approach that can be mentioned. In the perspective of the World War I centenary in 2014, the constitution of a virtual collection of family memory items illustrating the 1914-1918 conflict is currently under progress. The *Europeana 1914-1918* project<sup>11</sup> builds on the British *Great War Archive* initiative: it invites the European citizen to contribute family letters, objects, photos etc... during “collection days” that are organised at various locations throughout Europe. The items are examined and selected by experts, scanned and photographed, annotated with descriptive metadata. This public operation allows to collect not only historical items to build new collections, but also rich and authentic stories that go with them.

### ***Share and reuse***

Another trend regards the evolution of the Web and associated technologies. This technical evolution influences how digital libraries are built and how the digital collections and objects are perceived.

A digitized resource doesn't stand alone on its own on a web page or is not just part of a results list but it is part of a thematic story, a scientific analysis, an art exhibition. Connecting an object with some context – or making sure that this becomes a possibility during the object's life – is a user expectation that European practitioners take into account in the processes to create and upgrade digital content.

The Amsterdam Museum as Linked Open Data<sup>12</sup> is an example of a growing practice, in line with the development of the semantic web: the digital objects of 70.000 artifacts of the museum have been prepared in such way that they can be connected to other rich descriptions available on the web like dictionaries, corpuses of subjects, ontologies etc... In the other way around, the museum's digital data are now also available to be connected to, from external users working on other online services and wishing to reuse them.

Making sure that a digital object can be re-used outside of its hosting institution is part of many implementation plans in Europe. Nowadays, more and more web content is created by end users by mixing and mashing-up various pieces of text, images and videos. Therefore the possibility to export and embed digitized content becomes a systematic business requirement. We can cite for instance the videos of the Institut National de l'Audiovisuel [<http://www.ina.fr>] – the French audio-visual archive – that can be shared via email and social networks through permanent URL, or that can be embedded in external websites like YouTube videos.

---

11 <http://www.europeana1914-1918.eu/>

12 <http://thedatahub.org/dataset/amsterdam-museum-as-edm-lod>