

Dig *Italia*

Anno IV, Numero 1 - **2009**

Rivista del digitale nei beni culturali

ICCU-ROMA

Think culture: Europeana.eu from concept to construction

Jon Purday
Europeana

Introduction

Europeana.eu, Europe's digital library, museum and archive, was launched on 20 November 2008 by the President of the European Commission, José Manuel Barroso. The portal gave access to 4.5 million digital objects from over 1.000 contributing heritage collections from every member of the European Union via a multilingual interface. Overwhelmed by public interest – some 10 million hits an hour – the site slowed to a crawl and was taken down for a substantial hardware reconfiguration. Back up in December in test mode with limited interactive features, the full rebuild and stress testing were completed within the first quarter of 2009, and all elements of the service were fully functioning by April 2009. This article looks at the genesis, development, launch and future of Europeana.

Political endorsement

The catalyst for Europeana was a letter sent by Jacques Chirac, President of France, together with the premiers of Germany, Spain, Italy, Poland and Hungary to President Barroso in April 2005. The letter recommends the creation of a virtual European library, to make Europe's cultural heritage accessible for all. The first paragraph carries the weight of the letter:

«Le patrimoine des bibliothèques européennes est d'une richesse et d'une diversité sans égales. Il exprime l'universalisme d'un continent qui, tout au long de son histoire, a dialogué avec la reste du monde. Pourtant, s'il n'est pas numérisé et rendu accessible en ligne, ce patrimoine pourrait, demain, ne pas occuper toute sa place dans la future géographie des savoirs»¹.

[The heritage held in Europe's libraries is of unequalled richness and diversity. It exemplifies the universal outlook of a continent which, throughout its long history, has been in constant dialogue with the wider world. However, if it is not digitised and made accessible online, this heritage will not occupy its rightful place in the future knowledge landscape.]

The catalyst for this letter had been Google's announcement of a programme to digitise the printed word, and their partnerships with a range of major US and UK libraries. There was concern in Europe that the project would be directed towards

¹ *Letter of 28 April 2005*, http://ec.europa.eu/information_society/activities/digital_libraries/doc/letter_1/index_en.htm.

Anglophone content and that a complementary approach should be taken to digitise the European intellectual tradition in its original languages.

There was also disquiet that the large scale digitisation plans of Google and Microsoft would transfer a significant amount of public domain intellectual resource into the private sector, and that therefore equivalent European programmes ought to be conceived of as broadly open access, open source and non-exclusive.

The letter added resonance to the work that the Commission's Information Society and Media Directorate had been engaged in for over a decade, with programmes such as Telematics for Libraries. It gave strong political endorsement to the Directorate's strategy, *i2010: communication on digital libraries*² which was published on 30 September 2005. The strategy announced the intention to promote and support the creation of a European Digital Library, as a goal within the European Information Society *i2010 Initiative*³, which aims to foster growth in the information society and media industries.

The intention was endorsed by the Council of Ministers in November 2006, followed in September 2007 by a vote in the European Parliament which overwhelmingly adopted the Commission plan.

From the first, the purpose of the Digital Libraries Initiative was «to make Europe's cultural, audiovisual and scientific heritage accessible to all»⁴. The scope was thus far wider than simply a library digitisation project, and the target audience conceived in very broad terms.

The concept was to create a space in which all manifestations of Europe's cultural and scientific heritage could be connected and integrated within a single portal, in a multilingual environment. In part, the idea took shape because technology now enabled it; to a greater degree, however, it was born of a sudden leap in user expectations. Anybody who was using Web 2.0 sites was used to being able to watch video, listen to audio, see images or read text in the same space.

The European heritage and information domains – museums, libraries, archives and audio-visual collections – had been digitising significant pictures, films, books, photographs, sounds, newspapers, manuscripts and archival records over the previous decade. While the digital material featured in the catalogues and

¹ *Letter of 28 April 2005*, http://ec.europa.eu/information_society/activities/digital_libraries/doc/letter_1/index_en.htm.

² *Communication from the Commission of 30 September 2005, "i2010: digital libraries"*, SEC(2005) 1194, SEC(2005) 1195, <http://europa.eu/scadplus/leg/en/lvb/l24226i.htm>.

³ *Europe's Information Society, i2010 - A European Information Society for growth and employment*, http://ec.europa.eu/information_society/eeurope/i2010/index_en.htm.

⁴ *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, i2010: Digital libraries*, SEC(2005) 1194, SEC(2005) 1195, <http://europa.eu/scadplus/leg/en/lvb/l24226i.htm>.

databases of individual institutions, it was deep Web content and particular items were rarely findable by search engines. It was recognized that what users couldn't find online didn't exist for them. Furthermore, users didn't want, or weren't able, to make judgments about where an item of interest was held – either which country or what type of organization.

Heritage organisations recognized that Europeana offered an opportunity to renew their relevance to a new generation of Web users and meet changing expectations. To give search results that included relevant manuscript material, books by and about a subject, original newspaper reports and photos of the subject, TV coverage and film footage, was to enrich the users' experience. Users were also more likely to trust the provenance of material from named heritage organizations, and such authenticity added value to the content.

The Commission also recognized the potential economic benefit of opening up access to public domain cultural content. Free access to a critical mass of content, with possibilities for re-use, would stimulate new ideas and opportunities for adding value within creative enterprises, and generate new services in the education and tourism sectors.

Project antecedents

Development of a European Digital Library, focused exclusively on digitised books, began in France in May 2006, when the French government invited the Bibliothèque nationale de France (BnF) to produce a proof of concept in the national library arena. The BnF began a project with Portugal and Hungary to integrate digitised texts and the result was previewed in October 2006. After its acceptance a demonstrator was built and released in March 2007⁵. It included about 7.000 books from the BnF, 4.000 from the National Library of Hungary, and 1.000 from the National Library of Portugal.

The Conference for European National Librarians (CENL) was also asked in November 2006 to submit a proposal for the European Digital Library Project which would include a roadmap for a cross-domain portal. CENL had developed and launched the European Library, a successful service that integrated the catalogues of all the national libraries of 46 European countries. The experience of pan-European working, the international team, the proven infrastructure, and the base in the Koninklijke Bibliotheek – the national library of the Netherlands – in The Hague, gave it firm foundations on which to develop the European digital portal project.

⁵ *Réunion de présentation du projet de Bibliothèque numérique européenne (BNUE) au Président de la République, M. Jacques CHIRAC*, http://www.elysee.fr/elysee/elysee.fr/francais_archives/actualites/a_elysee/2007/mars/reunion_de_presentation_de_la_bibliotheque_numerique_europeenne-bnue.74358.html.

The Thematic Network: EDLnet

The project that would begin the building of Europeana was funded by the Commission under their eContentplus programme, one of the research and development funding streams of i2010. It was called the European Digital Library network (EDLnet) and funded as a Thematic Network⁶, with € 1.3 million for the 18-month project coming from the Commission. Overheads in the form of services and accommodation costs (equivalent to some 20% of the funding) were covered by the project’s host, the Koninklijke Bibliotheek. The Dutch Ministry of Education, Culture and Science (Ministerie van Onderwijs, Cultuur en Wetenschap, OCW) also contributed to significant marketing and communications costs which were outside the scope of the funding agreement.

Professional associations	
ACE	Association Cinémathèques Européennes
CENL	Conference of European National Librarians
CERL	Consortium of European Research Libraries
EMF	European Museum Forum
EURBICA	European Regional Branch of the International Council on Archives
FIAT	International Federation of Television Archives
IASA	International Association of Sound and Audiovisual Archives
ICOM Europe	International Council of Museums, Europe
LIBER	Ligue des Bibliothèques Européennes de Recherche
MICHAEL	Multilingual Inventory of Cultural Heritage in Europe
Institutions and ministries	
KB	Koninklijke Bibliotheek, the National Library of the Netherlands
	Bundesarchiv, Germany
BnF	Bibliothèque nationale de France
INA	Institut National de l’Audiovisuel, France
BNE	Biblioteca Nacional de España
MCU	Ministerio de Cultura – España

Figure 1. The board of participants, EDL Foundation, at the close of the EDLnet project in March 2009. Subsequently, membership has increased and a new governance structure implemented.

⁶ The Commission funds Thematic Networks, providing 100% of direct costs and not including overheads, and Best Practice Networks, providing 80% of direct costs and not including overheads.

The project's first task was to create a legal entity – the EDL Foundation – to own and carry forward the development of the service, bid for further funding and employ staff. Cross-domain interoperability was a founding principle: the idea that collaboration and knowledge transfer between the museums, libraries, archives and audiovisual collections was fundamental to the success of the enterprise. Accordingly, the Foundation embedded the concept at its core, bringing together many of the key professional associations on its Board – see Figure 1.

Engaging the international associations of each of the major domains at Board level endorsed the Europeana project and promoted the movement towards interoperability at the highest professional level.

Group Characteristics

Five user profiles for end users of the Europeana service were identified:

1. General User
2. School student
3. Academic user (undergraduates, postgraduates and teachers)
4. Expert Researcher
5. Professional user, e.g. librarian, archivist, curator

These user groups were characterised as follows:

1. General users have a generic interest in culture or history. They are familiar with basic search functionalities, have no specific domain knowledge, are “google-minded” and visit sites that have large volumes of content to offer, such as YouTube and Wikipedia.
2. School students will make use of the service as part of educational exercises. Culture and heritage are incorporated in many school curricula, which means that Europeana could be used in a variety of educational contexts. School students will expect the service to be easily accessible, immediately appealing, visually attractive or even playful, easy to handle while dealing with their projects, and without jargon.
3. Academic users represent the other end of the educational spectrum. They may have excellent domain knowledge, or aspire to achieve that. They will expect the information offered to be comprehensive, accurate, representative if not complete, and easy to reuse in the context of educational assignments.
4. Expert researchers look for specific information on a specific topic. They are to a certain degree skilled in using retrieval services and may make use of the advanced search button to get the most out of the system. As this group is most likely to publish the results of the research, it includes users who are

prepared to buy something or travel to visit the contributing institutions.

5. Professional users are most likely staff members of a cultural heritage organisation. They are skilled in using information systems, but with a different perspective than expert researchers. They may be interested in details as well as very generic information, for instance for improving information services of their own institution.

Looking at the motives of these user groups in relation to Europeana, it is possible to identify four types of objectives:

1. The user wants to be entertained. This includes users who have time available to browse around the Internet and have a structural or incidental interest in cultural heritage. They come to Europeana because they expect there to be a lot of interesting content. For these users it is not important what they find is as long as it is engaging.
2. The user wants to know more about a cultural or historic subject or person. This includes users that have a specific reason for their interest: it could be that they need to do a project for school, study or work on a certain subject, or that they have been made aware of a certain subject through current news or in a conversation with colleagues, friends or family. These users are looking for the most relevant results and would not want to see lots of results that are not relevant to them. To be able to determine what is relevant to them, information about the specific objective of the user is necessary.
3. The user wants to know the current whereabouts of cultural heritage materials. This includes users that are planning to see the original objects for research purposes, or users that are about to undertake a trip and would like to know what cultural heritage they can visit during a holiday or other type of stay. These people will also be interested in getting more information on interesting events and collections in the area, as well as local services such as guided tours.
4. The user wants to be part of a community of interest. This includes users who may be students, researchers or members of a cultural society and want to share their knowledge via an online environment such as a social platform with a cultural focus. They may want to present their opinions and ratings of cultural heritage resources to their network of family, friends and colleagues as well as share personal items (photographs, documents etc.)

Figure 2. User profiles and scenarios

It also enabled the four domains to communicate the objectives widely to their networks and members, and encourage widespread and diverse contributions of content.

The scope of the Foundation's membership enabled the European Digital Library network – EDLnet – to recruit a network of over 100 heritage organisations, universities, research institutes and related technical projects. They represented every European member state, and every heritage domain. Their representatives gathered for the first plenary session of the project at the Koninklijke Bibliotheek in September 2007.

It was recognised from the first, by both the partners and the Commission, that the name European Digital Library created a set of expectations among end-users, who were likely to equate it only with digitised books, rather than the wider range of cultural heritage resources. Equally, representatives from across the domains considered that the use of the term “library” appeared to privilege one group of content providers above others. After wide consultation with stakeholders and users, Europeana was selected as a name that was widely understood.

The name Europeana had originally been used for the BnF's pilot, and the BnF generously offered the trademark and domain name Europeana.eu to the project. The Latin word for Europe, Europeana also carries the meaning of “things European”. It needed no translation, and expressed in essence what users could find at the site.

The remit of the new project was to build a proof of concept prototype of a cross-border, cross-domain, user-centred service. This had to be ready for launch by November 2008, and provide access to a minimum of two million digital items.

The task was split into work packages, each involving around 15 key people, selected from among the partners for their particular expertise. Key work packages included:

- Work package 1: Human, Political & Intercommunity Interoperability. Led by David Dawson, then of the Museums, Library and Archives Council, UK, and Mats Lindquist from ABM Centrum in Sweden. Tasks included determining the best organisational model for an operational Europeana portal from a stakeholder or content providers' viewpoint. Outputs included the final draft of a sustainable business model for Europeana and the roadmap to achieve it.
- Work package 2.1: Standards and Interoperability of Standards (focused on metadata). Led by Makx Dekkers, an independent consultant, this work group agreed common metadata standards and models in relation to the operational models (centralised and distributed). It identified potential common sets of vocabularies and classification systems for various descriptive properties (object types, subjects, audiences, rights etc).
- Work package 2.2: Semantic and Multilingual Interoperability. Led by Stefan

Gradmann of Humboldt University, Berlin. This work package selected a set of consistent object models against the initial specification of user requirements and the identification of the domains to be integrated. It selected service description frameworks and identified the extent to which translation tools (metadata, object data) and multilingual vocabularies could assist in multilingual access to Europeana.

- Work package 2.3: Technical Architecture. Led by Carlo Meghini, of Consiglio nazionale delle ricerche, Istituto di scienza e tecnologie dell'informazione. This work package selected the suite of protocols and formats used to support the organisational and semantic interoperability and integration with other data collections. It also looked at solutions for accessibility, security, authentication, integrity and made recommendations on required network bandwidth.
- Work package 3: Users and Usability. Led by Daniel Teruggi from the Institut National d'Audiovisuel, the package created initial user requirements for the features and functionality required in the portal. It clarified the shared requirements across the four domains and identified those specific to each. It created a common set of recommendations by reducing the differences and extending areas of agreement on users and usability.

Defining the user requirements

Work package 3 brought together webmasters and Web marketing experts from each of the heritage domains. Three workshops were held for groups of invited experts in The Hague, Paris and London. From these sessions came the definitions of target users of Europeana, and their potential scenarios of use.

With these scenarios in place, the workshops were also able to map out the use cases that shaped Europeana in terms of functionality: searching, refining, browsing, saving, personalising, tagging, sharing and community building.

One clear requirement coming from the workshops was the need for the site to be available in all the 23 official languages of the EU. This had been identified early in the Commission's thinking. While multilingual search was ruled out as too complex for the prototype to implement in the short time available, it was practical to make all the pages of the prototype's interface available in all 23 languages, plus Catalan and Icelandic.

Public previews of the portal

At two stages in the development of Europeana, focus groups with end users were held to validate these use cases in relation to what had been developed.

The first of these key milestones along the route to Europeana was in February 2008 with the creation of the *maquette* – the first demo version – that gave partners, content contributors and users a sense of the look, feel and functionality of the site. This was shown in the form of a Flash animation, with dummy content in place, at a plenary conference in Frankfurt in February 2008.

Structured feedback from the 200 delegates was recorded, then focus groups with end-users and librarians who had no prior knowledge of Europeana were held in Amsterdam, Stockholm and Warwick. This qualitative analysis was then amplified by a quantitative online survey which garnered 700 responses. Adjustments were made to the site's navigation and look and feel in response to the findings of the research.

As the project began to build an initial live version of the site, changes were also made to the interface and functionality because of issues deriving from the content itself. A map showing location of items was shown in the *maquette*, for example, but was omitted from later releases because the lack of geographic co-ordinates in the metadata made it too complex to implement in the prototype stage of Europeana.

The second and final public demo version, this time with real content in the database and searched in real time, was shown at the project conference in June 2008 in The Hague. Again, 150 delegates spent time reporting back on explorations of the live demo version, and focus groups with end-users were also held in Paris, Athens and Bucharest.

At this stage the functionality and navigation of the site were endorsed, but it was felt that the look and feel, which had worked well for the project development, would need to be refreshed. This led to a fundamental reappraisal of the branding and the appointment of the SalterBaxter agency in London to devise a new approach. Their resulting visuals were circulated to members of the users and usability work package who had volunteered to provide detailed feedback, and a new brand concept with a strong visual signature was chosen. It was engaging, reflected a diversity of content in a contemporary framework, and centred on the search box.

Content collection and ingestion

In order to begin the process of content acquisition, a questionnaire was put on the project development website early, following the Frankfurt conference in February 2008. Links were sent out to all partners and more widely to their national, professional and membership networks. The questionnaire sought information about subject matter, quantity, file format and metadata standards. 85 responses were received, recording substantial numbers of digitised items that organisations were prepared to make available to Europeana. Pan-European coverage was good, but the same was not true across domains: the majority of respondents were libraries, and audiovisual collections were not strongly represented. There were, however, a number of important national aggregators responding, which represent a broad spectrum of European organisations. Because national aggregators work across the domains, they have already dealt with homogenisation of metadata from their contributing institutions. The largest aggregator was

Culture.fr⁷, which accounted for two million items from 480 libraries, museums and archives in France. Bringing in material for aggregators became a top priority, because by doing so Europeana fulfilled the cross-domain remit, and gained more experience of finding solutions to interoperability problems by working with technical colleagues from aggregator sites.

It was also important to represent every European member state and to include a rich cross-section of formats and carriers. Digitised materials offered to Europeana included digitised books, manuscripts, paintings, engravings, drawings, archival documents, photographs, posters, postcards, newspapers, maps, film footage, TV broadcasts, field recordings, discs, wax cylinders, museum objects and specimens.

Much of the harvesting and data preparation for the Europeana prototype was done manually. Initially the Europeana team worked closely with the content provider to specify the required elements. The datasets, expressed in XML, were then harvested. Approximately half the harvesting was via Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH)⁸, and half via ftp servers. A small percentage was delivered as email attachments, and under 1% on DVD.

The harvested data was cleaned and pre-processed. All characters are made XML compliant and encoded in UTF8, the Unicode 8-bit encoding format that ensures compatibility between systems. Europeana uses UTF8 to encode all character sets, including Greek and Cyrillic.

Sample data was then analysed to show the content of all fields. This showed the layout for mapping purposes and enabled a check to ensure all mandatory fields were present. Once the mapping rules had been specified for each dataset – a fairly labour-intensive process in the prototype – the metadata was run through transforming tools to fit the Europeana data schema. In some cases, the content provider had already done the mapping; this proved valuable, and will be a requirement going forward.

During the mapping process, normalisation took place. For example, the date was expressed as a four-digit year, the country of provision was made consistent, and the item allocated to one of four format types: image, text, sound and video. This broad grouping enabled users to refine searches by the material types that they readily identified with. Providers' names were also added to each record, and an Authority File of providers maintained. The data was then fed into the indexing engine, SOLR Lucene, which indexes on all fields.

The metadata contains two or three URIs (Uniform Resource Identifiers):

- one that points to a thumbnail;

⁷ Culture.fr, <http://www.culture.fr>.

⁸ Open Archives Initiative Protocol for Metadata Harvesting, <http://www.openarchives.org/pmh>.

- one that points to a larger scale file;
- one that points to the item within its Web page on the provider's site.

Of the URIs, 1 is the required whenever possible. If 1 cannot be provided, 2 can be used to generate the thumbnail. Europeana must have either 2 or 3, but preferably both. URIs 2 and 3 must be persistently identifiable.

The URIs relating to the thumbnails are used to cache them as references on the image server. This is done to speed up the retrieval and display of results. Search results are displayed 12 to a page, showing a thumbnail of the title page, image or still plus title/creator/provider metadata. If the thumbnails were all retrieved or generated on the fly from 12 different host sites this would delay the retrieval process for the user – hence the importance of caching, so that results appear instantly to the user. In cases where a book is digitised as a PDF, Europeana uses a tool that takes the title page as the thumbnail. Sound files are represented by a default sound icon.

One problem that has been experienced is the thinness of the metadata. The metadata is often very basic, in some cases because digitisation programmes have put their resource into the digitisation of items rather than their cataloguing. The lack of rich metadata has an impact on how accurately search results are retrieved, and how material displays in the timeline. It is also the case that in order to be able to incorporate the widest variety of standards, Europeana has adopted a limited metadata schema, which is not always able to reflect the richness of some of the metadata that is supplied.

This lowest common denominator approach was necessitated by the diversity of metadata standards that are used in different institutions, different domains, different countries.

The experience gained from harvesting and ingesting the first major tranche of metadata, representing the 4.5 million items that formed the content of the public prototype, enabled the Europeana team to produce the Specification for the ESE (Europeana Semantic Elements)⁹.

The European Semantic Elements are DC (Dublin Core) qualified. They comprise a subset of the DC terms¹⁰ and a set of twelve elements which were created to meet Europeana's needs. In future, rather than map manually within the Europeana office, all content providers will need to map to the ESE prior to data harvesting.

Note that the DC and Europeana namespaces both have "Type" and "Language" elements. These have different uses which are explained in the full specifications online.

⁹ *Specification for the Europeana Semantic Elements, Version 3.1*, 25 February 2009, http://version1.europeana.eu/c/document_library/get_file?uuid=5fdb5df9-bde5-42c4-8df5-7c9e488e7cf7&groupId=10128.

¹⁰ DCMI Usage Board, *DCMI Metadata Terms*, 14 January 2008, <http://dublincore.org/documents/dcmi-terms>.

	Libraries	Museums	Archives	Audiovisual
Dublin Core	X	X	X	X
TEL Application Profile	X			
MARC (21 etc)	X			
FRBR	X			
EAD			X	
CIDOC-CRM		X		
MODS	X			
Museumdat		X		
EBU Core				X
VRA		X		
CDWALite		X		
MPEG-7				X

Figure 3. Standards used by content providers

Object surrogacy

One of the important decisions taken early in Europeana’s development was to create a surrogate of the digital object that would comprise, in its simplest form, a set of metadata, a small image or thumbnail of the digital object (except for sound which would be represented by a sound icon) and a URI, a persistent identifier that would link to the full resolution digital object in its own website.

Within Europeana, users would explore the surrogate only. In order to see the full digital object, they would be given a link back to the digital object in the provider’s website.

This use of the surrogate had several advantages:

- it meant that the tools associated with playing, page turning, magnification and exploration of the full object would be used at the content provider’s site, obviating the necessity for Europeana to deal with every single file and player format;
- massive servers to hold duplicate content from the heritage organisations of Europe would not be needed;
- digital updating and preservation issues remained the responsibility of the content provider, not of Europeana;
- Intellectual Property Rights (IPR) in the high resolution file remained squarely within the control of the content provider;
- it enabled the content provider’s identity and branding to be visible to the user, reinforcing provenance and authenticity.

The Europeana Semantic Elements are:

Source	Element	Element Refinement(s)
DC	title	alternative
DC	creator	
DC	subject	
DC	description	tableOfContents
DC	publisher	
DC	contributor	
DC	date	created; issued
DC	type	
DC	format	extent; medium
DC	identifier	
DC	source	
DC	language	
DC	relation	isVersionOf; hasVersion; isReplacedBy; replaces; isRequiredBy; requires; isPartOf; hasPart; isReferencedBy; references; isFormatOf; hasFormat;
Europeana		conformsTo isShownBy; isShownAt
DC	coverage	spatial; temporal
DC	rights	
DC terms	provenance	
Europeana	userTag	
Europeana	unstored	
Europeana	object	
Europeana	language	
Europeana	provider	
Europeana	type	
Europeana	uri	
Europeana	uri	
Europeana	year	
Europeana	hasObject	
Europeana	country	

Figure 4. Europeana Semantic Elements (ESE)

System usage

The potential level of visits to the prototype site at its launch were assessed by benchmarking against the BnF and the British Library (BL). They were consulted because both had held high-visibility launches of cultural content sites. The resulting figures were then extrapolated to cover a pan-European audience, i.e. some ten times greater than that reached by the BnF or the BL.

A cost-effective environment was deployed with four virtual servers – two front end, one search server and one database server. The configuration was stress-tested to take a maximum load of 3.000 concurrent users and 5 million hits per hour.

Launching Europeana

The launch of the prototype of Europeana.eu was scheduled for 20th November 2008. It was preceded by a press briefing on 19th organised by the Commission's press office in the main press auditorium in Berlaymont, Brussels. It featured a promotional video¹⁴, a scripted live demo of the website and an extended question and answer session. The press release was issued in the 23 official languages of the EU and was accompanied by raw footage of digitisation activities shot by the Commission's AV unit at the BnF, Beeld en Geluid [Sound and Vision archive of the Netherlands], and the Rijksmuseum, Amsterdam. This was for use in TV news packages as backdrop, enabling journalists to voiceover their report about the launch in their own language.

At its beta launch, the site gave access to 4.5 million digital objects – more than double the initial target – from over 1.000 contributing organisations, including world-famous national library, gallery and museum collections from the capitals of Europe. The stories that could be told using this content captured the imagination of the press, so that on the morning of the 20 November, Europeana was widely represented in the print, broadcast and online media. By midday it was already experiencing heavy traffic, estimated at around ten million hits an hour.

The formal launch took place at 14.00 in the Palais Charles de Lorraine in the Royal Library, Brussels. It followed a meeting of the Council of Ministers of culture and education from the member states. The audience of ministers, stakeholders, policy-makers and press was addressed by the President of the European Commission, José Manuel Barroso, by Mme Viviane Reding, European Commissioner for Information Society and Media, and by Mme Christine Albanel, French Minister of Culture and Communication and acting President of the EU Council of Culture Ministers during the French presidency.

José Manuel Barroso outlined the ambitious vision behind the Commission's support:

¹⁴ Europeana, *About us*, <http://www.europeana.eu/portal/aboutus.html>.

«I believe that Europeana has the potential to change the way people see European culture. It will make it easier for our citizens to appreciate their own past, but also to become more aware of their common European identity»¹⁵.

The broadcast news coverage of the launch event pushed usage up another notch and by late afternoon it was peaking at an estimated 13 million hits an hour. Instead of taking milliseconds to execute, as had been the case during stress testing, a search was taking 15 minutes to return results. It proved impossible to scale up the virtual configuration sufficiently to deal with the load: the site remained unresponsive and the user experience poor. In order to reconfigure the system, the Commission and Europeana's senior management reluctantly decided to take the site down late in the evening of the 20th November.

A message was put on the site in English, French and German, apologising that it had been taken down, and linking visitors back to the project development site. This site was quickly reworked to include explanations, the demonstration of the site seen at the press conference, the promotional video, the press releases and their associated Questions and Answers. This development site was then very heavily visited, as Europeana, a "victim of its own success" made headlines for a second day, and 80.000 people signed up for the eNews to keep in touch with developments.

The first steps in the rebuild were to change from a virtual environment to dedicated hardware for the servers and load balancers, and to increase the number of servers. The portal went live again in test mode in mid-December, still with the "beta" flash and also an explanation that the system was being reconfigured and that a limit would be placed on the number of visitors during this test phase. The interactive features of the site, including the carousel, the feature "people are currently thinking about" and "My Europeana", the personal user space, were temporarily disabled, because the new hardware required a different configuration and a software rewrite to manage them.

The rebuild put in place hardware load balancers, with dedicated image and indexing servers, so the system is now equipped to handle much higher concurrency than on first launch. The search engine, SOLR Lucene, is an open source, fast indexing server which has proved capable, in renewed and increasingly rigorous stress testing, of achieving state of the art search speeds.

The interactive features went live again in March 2009, and in April all testing had been completed, and Europeana returned to normal service.

¹⁵ José Manuel Durão Barroso, *Europeana: A shop window on Europe's cultural heritage*, SPEECH/08/632, Launch of Europeana, Brussels 20 November 2008, p. 2, <http://europa.eu/rapid/pressReleasesAction.do?reference=SPEECH/08/632&format=PDF&aged=0&language=EN&guiLanguage=en>, p. 2.

Europeana's future

The final round of eContentplus funding allocated € 69 million over 2009-2011 for research on digital libraries, and another € 50 million to help improve access to Europe's cultural and scientific heritage.

Europeana itself will receive € 2 million a year from the Commission in 2009-11, with Member States, cultural institutions and the private sector gradually increasing their share of funding as the project develops.

The EDL Foundation is a partner in a cluster of projects known as the Europeana Group – see Figure 6. The projects have been funded by eContentPlus to develop technology and services for Europeana, and to digitise and channel content into the portal.

The projects providing content to Europeana are mainly domain aggregators, meaning that they collect content from their heritage sector – museums, archives, libraries and audiovisual collections – and channel it into Europeana.

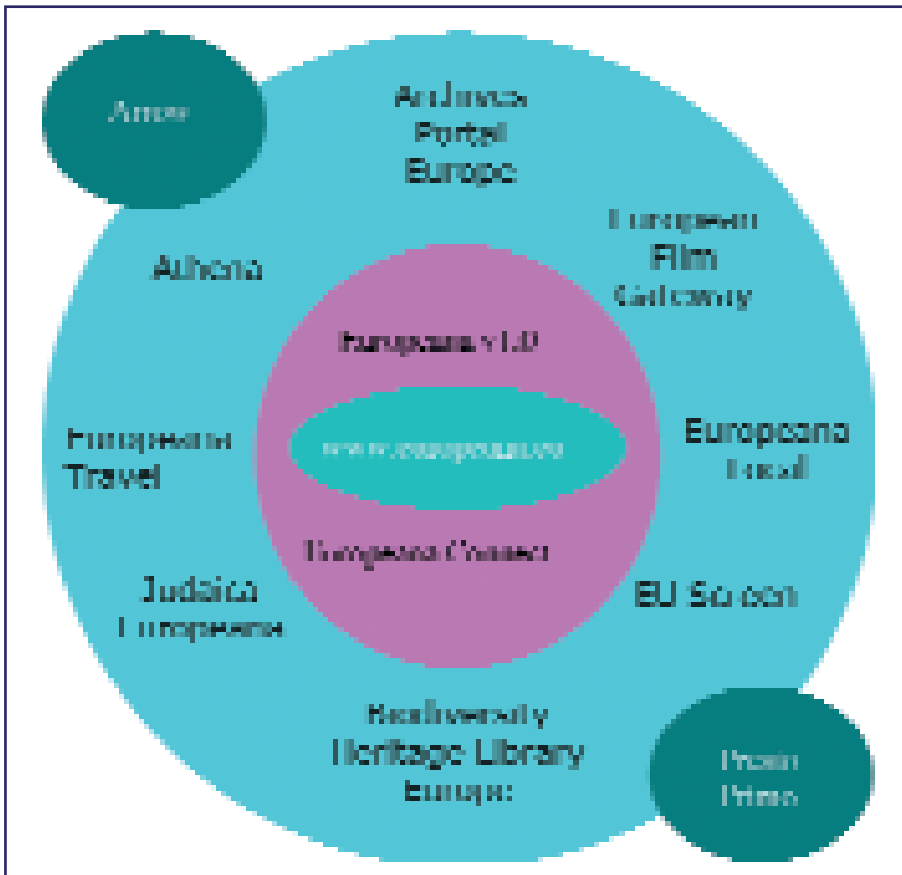


Figure 6. The Europeana Group of projects

In some cases, they will also create their own portal, for example European Film Gateway, which will not only show film content, but also a wider range of general information about European film archives.

If Europeana is to be scaleable in the future, and reach the projected targets of 10 million items in 2010, 25 million items by 2012, it will need to work with a limited number of these domain and national aggregators¹⁶ rather than with a plethora of individual institutions. These aggregators will deal with issues around harvesting, mapping, ingestion and IPR for their domain and channel the standardised meta-data into Europeana.

The Europeana group: core projects

Europeana version 1.0

The key project supported by the EDL Foundation is Europeana version 1.0. This 30-month project began in February 2009 and will develop the prototype into a fully-operational service. It is 100% funded, excluding overheads, with the Koninklijke Bibliotheek providing services and accommodation.

Retaining broadly the same partners as EDLnet, Europeana v1.0 is again a Thematic Network¹⁷. In outline, its mandate is to solve key operational issues related to the implementation and functioning of Europeana.

These include:

- Developing partnerships to ensure a rich content flow from national and domain aggregators:
 - creation of licence and partnership agreements;
 - rights management and IPR;
 - management of partner content and continuous access to that content;
 - promotion of standards and data models.
- Building the back end systems needed to manage the delivery and access of this content:
 - technical building of the operational service – functional specification implementation;
 - automated content ingestion workflow;
 - implementation of new data and object models;
 - service agreements with bodies involved in running the service.

¹⁶ National aggregators include, for example, Culture.fr which works with hundreds of museums, libraries and archives throughout France. Culture.fr integrates all types of content and successfully submitted very large quantities to the Europeana prototype.

¹⁷ The Commission funds Thematic Networks, providing 100% of direct costs and not including overheads, and Best Practice Networks, providing 80% of direct costs and not including overheads.

- Managing the channels enabling other environments to use the content made interoperable by Europeana, via Web services or APIs (Application Programming Interfaces):
 - distribution to other channels for mash-ups and re-use;
 - implementation of new plug-ins and technologies.

Europeana Connect

Connect is closely associated with v1.0 and is a 30-month Best Practice Network that will undertake the specific technical implementation work, to increase the amount of interoperable content that can be accessed in Europeana. Co-ordinated by the Austrian National Library, Europeana Connect will deliver core components and value-added services for Europeana. It will facilitate multilingual access by implementing a suite of translation tools and language resources. This will take the project a step closer to the multilingual grail: that users should be able to input search terms in their own language, which are translated on the fly during the search process, and have hits returned in a range of selected languages.

Connect will also build a layer of semantic data which will be the basis for all semantic processing in Europeana. Semantic processing would, for example, enable the relationship between “the Blessed Virgin Mary”, “the Madonna”, “the Mother of Christ” and “the Holy Virgin” to be recognised during the search process. So a user entering just one of these terms would see also results for the linked concepts. Connect will provide the technologies and resources to semantically enrich vast amounts of digital content in Europeana.

This will enable semantically-based content discovery including support for advanced searching and browsing, allow for delivery of enhanced services and make Europeana content more accessible, reusable and exploitable.

Other deliverables from Connect will add value to European’s services, for example:

- A map linked to a more sophisticated timeline tool that will allow users to browse and search in time and space
- An interface for mobile devices and implement the middleware for adapting Europeana content for mobile usage and presentation
- An eBooks-on-demand service
- Multimedia annotation

Finally, Europeana Connect will add a critical mass of 200.000 audio tracks to Europeana, building an audio-aggregation infrastructure to harvest audio using OAI/PMH from some 150 audio archives.

The Europeana group: content providers

*Archives Portal Europe (APENet)*¹⁸

The portal will be the central reference point for the national archives of Europe, and will also channel archival resources into Europeana. APENet is a three-year project co-ordinated by the Spanish national archives which will bring together 16 million multilevel descriptions of archival fonds and documents, together with 31 million digitised archival objects.

*Athena (Access to cultural heritage networks across Europe)*¹⁹

An aggregator for museum content, especially at the national level, Athena will begin channelling content into Europeana from September 2009. It is a 30-month project that began in November 2008, coordinated by the Italian Ministry of Culture. Athena is setting up a network of national representatives to raise awareness of standards, develop technical architecture and provide tools to support content providers.

*BHL Europe (Biodiversity Heritage Library Europe)*²⁰

Coordinated by the Museum für Naturkunde, Berlin, the three-year BHL project began in May 2009. It will bring together museums, botanical gardens and other natural history collections to provide a multilingual access point for taxonomic materials and other biodiversity resources.

EU Screen

Beginning in October 2009, EU Screen will focus on television collections, digitising 30.000 programmes, stills and associated documentation. It will promote the harmonisation of access to Europe's TV heritage, evaluate the use of archival broadcasts in leisure, learning and research environments and produce guidelines for long-term rights clearance.

*EFG (European Film Gateway)*²¹

A three-year project that began in September 2008, EFG is coordinated by the Deutsches Filminstitut. It will create access to 700.000 items, including 24.000 hours of film, plus posters, scores, scripts and stills. It will devise agreed common interoperability standards for films, and establish best practice for clearance of rights.

*Europeana Local*²²

Focused on regional libraries, museums and archives, Europeana Local began in

¹⁸ <http://www.apenet.eu>.

¹⁹ <http://www.athenaeurope.org>.

²⁰ <http://www.biodiversitylibrary.org>.

²¹ <http://www.europeanfilmgateway.eu>.

²² <http://www.europeanalocal.eu>.

June 2008 and runs for three years. It will provide 20 million digital items over the course of the project, but will also enable all type of smaller organisations to contribute to Europeana in the long-term by promulgating standards and setting up easy processes for delivering content.

*Europeana Travel*²³

Bringing together the Ligue des Bibliothèques Européennes de Recherche (LIBER) and the Conference of European National Librarians, the two-year project began in May 2009 and is coordinated by the National Library of Estonia. It will digitise some 4.000 maps, 16.000 images, and 20.000 texts associated with travel, trade, tourism and migration. The project will establish and share best practice digitisation techniques throughout the extensive network of national and research libraries in Europe.

Content Enablers

A number of projects are facilitating work being done to bring content into Europeana.

These include:

*Arrow (Accessible Registries of Rights Information and Orphan Works)*²⁴

Is a consortium of European national libraries, publishers and collective management organisations. Arrow supports the i2010 Digital Library Project by finding ways to identify rightholders, rights and clarify the rights status of a work, including whether it is orphan or out of print. This will enable libraries as well as other users to obtain information on who are the rightholders, which are the relevant rights concerned, who owns and administers them and how and where they can seek permission to digitise and/or make available the work to user groups.

*PrestoPrime*²⁵

This project will research and develop practical solutions for the long-term preservation of digital media objects, programmes and collections, and find ways to increase access by integrating the media archives within Europeana and other digital libraries in a digital preservation framework. This will result in a range of tools and services, delivered through a networked Competence Centre. The Competence Centre and the European Association for Audiovisual Archives will be established to provide business models, registry and best practice services and training.

²³ <http://www.europeanatravel.eu>.

²⁴ <http://www.arrow-net.eu/about-arrow>.

²⁵ <http://www.prestoprime.org>.

Conclusion

Interest in Europeana remains high – at the political and policy-making level, from professional stakeholders and from the wider public. The project is funded well into 2011, and is developing a sustainable business model so that the service can develop in the medium to long term. Scalability will be proved over the next two years as the critical mass of content is provided by the Europeana group of projects. Beyond that, mass digitisation will drive Europeana's ingestion programme for the longer term.

The application of metadata standards across the domains is a critical success factor; related to this is the need for richer metadata. This can be achieved at one level by the institutions themselves; however, a greater benefit across the whole cultural heritage sector will be derived from the experiments in semantic enrichment that Europeana will take forward.

Clearly, users expressed a strong interest in Europeana when it was launched, but a key question is the extent to which this interest can be transformed into engagement. Success will be measured by how easily Europeana is able to put content into the users' workflow, by developing APIs and enabling mobile access. It will also need to offer a more participatory experience, enabling the users to generate content, to annotate, comment, mash-up and otherwise re-use the European cultural heritage to generate ideas and provide the inspiration for new cultural manifestations.

REFERENCES

- Robina Clayphan – Makx Dekkers – Kate Fernie – Stefan Gradmann – Catherine Lupovici – Carlo Meghini – Go Sugimoto – Julie Verleyen. *Specification for the Europeana Semantic Elements, Version 3.1*. 25 February 2009, http://dev.europeana.eu/public_documents/Specification_for_metadata_elements_in_the_Europeana_prototype.pdf.
- Makx Dekkers – Stefan Gradmann – Carlo Meghini. *Europeana Outline Functional Specification: For development of an operational European Digital Library, Version 1.7*. 1 March 2009, <http://europa.eu/rapid/pressReleasesAction.do?reference=SPEECH/08/632&format=PDF&aged=0&language=EN&guiLanguage=en>, p. 2.