



DELIVERABLE

Project Acronym: Ev3
Grant Agreement number: 620484
Project Title: Europeana Version 3

D5.3: Product Development Plan

Revision	Version 1
Date of submission	2014-06-30
Author(s)	David Haskiya, Europeana Foundation
Dissemination Level	Public

Project co-funded by the European Commission within the ICT Policy Support Programme

REVISION HISTORY AND STATEMENT OF ORIGINALITY

Revision History

Revision No.	Date	Author	Organisation	Description
1	2014-06-20	David Haskiya	Europeana Foundation	

Statement of originality:

This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

1 Background

This document describes product development priorities and aspirations for the Europeana platform in 2014. It follows the same format as the plans written in Europeana 2.0. While this is Europeana 3.0 deliverable it he decentralized nature of product development at Europeana means that the plan includes also the product development efforts undertaken in other projects than in Europeana 3.0 proper.

Current core technology/infrastructure projects

- Europeana Creative
- Europeana Cloud
- Europeana Sounds
- Europeana Space

Current projects delivering technology

- Europeana Inside
- Europeana Awareness
- LoCloud

Current content projects

- APEx
- Europeana Fashion
- Digitised Manuscripts to Europeana (DM2E)
- EUScreenXL
- Ambrosia
- PartagePlus
- Europeana Photography

Some of these projects all deliver or plan to deliver content to Europeana that can be indexed and displayed to best effect only by Europeana making significant changes in our systems. This is especially so for APEx which is a major stakeholder in our efforts to develop support for search and display of hierarchically related objects.

The approach taken in this document is to outline the overall **product priorities**, then outline the **portfolio of products** maintained and developed, and then visualise the development activities in the form of a simple **product roadmap**.

2 Product Priorities

The product development priorities fall under four primary categories:

2.1 Priority 1: Stability, performance and small-scale incremental development

As the number of currently committed products and services grows, an increasing proportion of the time of the Product-related activities at Europeana will be spent on maintenance and operational activities. Already developed and implemented features need to be bug fixed and kept up to date in order to provide a quality service to users.

Like any infrastructure project, performance remains a key challenge for Europeana. Products and websites are expected to be highly available and to handle large numbers of concurrent users and applications. Services such as the API have an even more explicit mandate to remain always available, as a variety of third-party applications and services are dependent on the Europeana API service. Other key production issues include the provisioning and rate limiting of API clients, scalability to meet the needs of search engine indexing spiders and stability of system-assigned uniform resource identifiers (URIs) for records in Europeana.

2.2 Priority 2: Improving data quality

As Europeana continues to expand its capacities to distribute data and content via new channels, it is increasingly clear that our audiences will expect clear and high-quality data to be made available to them. Existing problems with data generally originate at source, that is, in the documentation systems of data partners. While the products and systems of Europeana cannot directly address all of these quality issues we will focus on developing tools for quality assurance, data cleaning, data normalisation and data enrichment.

At the same time, data partners are increasingly asking for more transparency into the processes of ingestion and aggregation within Europeana. Providing some basic visibility into this process should facilitate closer relationships with the aggregation network, as well as make it easier to decentralise some of the Quality Assurance process. In short, to facilitate improved quality of incoming data and mappings to EDM.

Results of requirements in this area will perhaps initially be most visible to the data providers of Europeana, and to projects doing work in related areas. The end-goal of the priority is to improve the quality of our own end-user facing products and end-user facing products developed by others based on our APIs.

2.3 Priority 3: Supporting re-use and platformization

The next significant shift for Europeana as a whole is to become a platform that can enable an expanding sector of creative industries and others to make literal or transformative use of the contents of the Europeana repository. This will require an ongoing rebalancing of resources from the end-user search portal to the enabling technologies and APIs that will lower barriers to re-use.

Requirements developed under this priority will primarily be visible in the form of Europeana Labs and the APIs it exposed. This priority will also lead Europeana products to put a higher emphasis

on collections which encourage significant forms of reuse through the provision of sufficiently open licenses and direct links to content media of sufficiently high quality.

To better communicate Europeana as a unified platform, and to improve user satisfaction, a consistent Web Style Guide will be developed and implemented across all Europeana websites.

2.4 Priority 4: Supporting a Networked organisation

As Europeana is now at the centre of a complex and ever-growing ecosystem of projects, aggregators and networks, we need to be able to support efficient sharing of news, information and knowledge within the network. To that end we will during 2014 completely re-develop Europeana Professional (Pro) with a focus on improving its usability and stability based on extensive research and design processes.

As part of this development we will also improve our document management system with the goal of seamless integration with Pro in terms of improved on-site search and ease of upload.

3 Product portfolio and components

3.1 Overview

The core products of Europeana support the aggregation, enrichment, storage, index and distribution of scientific and cultural metadata aggregated from professional memory institutions. Some products also support the upload and display of user-contributed metadata and/or content, while some serve more of a documentary or marketing function, describing, contextualising or promoting the activities of Europeana and its partners.

The products currently being maintained by Europeana can be roughly classified into front-end and infrastructure products, and further classified based on the intended audience for each product. Of course, project Descriptions of Work sometimes define these products differently or under other names, but for planning purposes, classifications of the Europeana product portfolio include the following types:

- Front-end (generally Web) Products are intended for the direct use of end-users and are developed on either mobile or web platforms.
 - Business to Consumer (**B2C**). These are products and services targeted toward end-users who have direct and personal motivations for access. These could be more general motivations such as entertainment, or more specific or occupational motivations, such as pedagogic ones. B2C is also where Europeana reaches out to users indirectly via social media and specific partnerships like GLAMwiki/Wikimedia, Historypin, and Retronaut, and where Europeana directly measures impressions and viral reach.
 - This segment might be further subdivided to include Business to Researcher (**B2R**) products because it is assumed that researchers (university-level or higher) have unique needs in the search and display of cultural and scientific metadata and content. This will be a key consideration for 2015 when *Europeana Research* enters an intense phase of design and development.
 - Business to Business (**B2B**). These are products expected to be used people representing organisations, rather than accessed for reasons of personal interest or need. Because the needs of such organisations are quite diverse, in the Europeana context this is further broken down into two further segments.
 - Business to Network (**B2N**), whose services are primarily intended for Gallery, Library, Archive and Museum organisations (GLAMs), either now in the Europeana Network or prospective Network members.
 - Business to Developer (**B2D**). Services whose primary intended users are software developers and web developers.
- Infrastructure Products are developed and maintained for internal use within Europeana as part of line-of-business applications, or to support other systems.
 - Line-of-Business (**LOB**) products. These are applications that support an internal Europeana function, such as customer relationship management or metadata ingestion.
- **Core** services. These are generally products whose function is to store, index or transform data or content from the Europeana repository or its related stores of user-contributed content and data. These core services are developed to be in common to many current or future sets of user-facing products.

3.2 B2C

3.2.1 Portal

With the shift from Portal to Platform the development focus will be on features that are based on API-improvements or features that are of use also in the API. We will invest in fundamental improvements of our search service with the end result being improved quality of search and in improved cross-language retrieval.

Portal development will also focus on better display the increasingly rich EDM metadata we receive from our partners. This includes related (hierarchical) objects, contextual data from controlled vocabularies and authority files, richly documented and multiple media objects, browse/display of agents and concepts and user created annotations.

In terms of look and feel, and in order to reinforce the concept of Europeana as a unified platform, a coherent web style will be applied across all our websites including the portal. The aim is also that this will improve the usability and accessibility of the portal.

3.2.2 Exhibitions

While we would like to update the exhibitions to run on the Omeka 2.x CMS (rather than the outdated Omeka 1.2 it currently runs on) we currently lack the development resources and so have had to defer it. Doing the upgrade would be a prerequisite to adding new features and adapting the exhibitions to the new web style.

As curated and contextualised content is one area of research in Channels the future of the exhibitions CMS will be covered there. While formally not product development it may be worth to mention that we will experiment with creating and publishing exhibitions on the Google Cultural Institute.

3.2.3 Blog

The blog will be adapted to the new web style and become visually integrated with the Portal.

3.2.4 Europeana 1914-1918

We have in the last 12 months invested into improving our already successful WWI-site. Version 3.0 featuring institutional content from Europe and beyond was released and we have made further smaller improvements in the first half of 2014.

In Europeana 3.0 we will focus on improving the visual design of the site and aligning it closer to the portal, making it easier for the public to share and annotate content and also research a solution for the digital preservation of the user uploaded content. The technical framework underpinning Europeana 1914-1918 will be updated and better documented so that other institutions can adopt it for crowdsourcing campaigns of their own.

3.2.5 Europeana 1989

We will use Europeana 1989 as one development case in developing roundtripping capability between Historypin and Europeana. Otherwise no development is planned for the product.

3.2.6 Europeana Open Culture app and MUSE framework

We currently have no firm plans to upgrade the Open Culture app but will encourage partners and third parties to adopt and contribute to the framework for their own purposes.

3.2.7 Channels

In brief the Channels concept is to create a platform for Europeana and Network partners to with low-effort create vertical thematic channels for discovery and engagement. In 2014 the concept will be further researched and requirements and design defined for implementation in 2015. A music themed channel will be developed within Europeana Sounds and a map themed channel within Europeana 3.0.

3.3 Business to Network products (B2N)

3.3.1 Europeana Professional

As part of our efforts to become a truly networked organisation where news and knowledge is easily shared we will redevelop Europeana Professional with improved design, navigation and usability of features. The new version will be integrated with our Customer Relationship Management System and Document Management System.

As part of its redevelopment Pro will be upgraded to the new web style guide.

3.3.2 GLAM-wiki toolset

Developed by Europeana on behalf of the Wikimedia community we will maintain the toolset in 2014. We will apply to the Wikimedia Foundation for funding and if we receive it we will develop a new and improved version that will simplify doing batch uploads further and also make it possible for GLAMs to download improvements and additions created by the Commons community.

3.4 Business to Developer Products (B2D)

3.4.1 APIs & Linked Data

During Europeana 3.0 we will make major investments into improving and extending our existing APIs and developing new ones for specific purposes like user annotations and OAI-PMH harvesting. The individually largest extension of our API service will be the ability for developers to filter results to only objects that have media files meeting certain technical quality criteria combined with copyright status. This has long been the top requirement from our community of API-users.

In the area of Linked Data we will upgrade our publically available SPARQL-interface at data.europeana.eu to ensure it is constantly kept in synch with the main database serving our portal and other APIs.

Developing our APIs is crucial in our transition from portal to platform and all APIs will be made discoverable on Europeana Labs where they will also be documented.

3.4.2 Europeana Labs

Developing APIs is not in and of itself sufficient to power the transition to platform. The APIs need to be documented and that documentation be made available in a developer-friendly fashion. Apps created must be featured, datasets particularly suitable for re-use surfaced, the physical Labs part of the Europeana Network be made findable and the incubation services Europeana and partners offers published. All this will be done by further developing the [Europeana Labs](#) portal for creative industries.

3.5 Line of Business Products (LOB)

3.5.1 Aggregation tools

Metadata and content aggregation infrastructure is a key technological asset allowing Europeana to continuously harvest metadata from our content partners. These tools need to be maintained to ensure its readiness to efficiently operate growing volumes of metadata and previously unhandled digital content as well as fundamentally upgraded into its next generation. Key requirements of the new system are a unified approach towards storing metadata during various ingestion phases, automatic flow of the system till the final stage of publishing the ingested metadata and content, and more transparency provided to provider, achieved through more self-service ingestion tools and better reporting.

Finally the aggregation tools need to be adapted to be compatible with the cloud-based storage system developed in Europeana Cloud.

3.5.2 Enrichment & Data Quality

Enriched, normalised and consistent quality metadata is not only a purpose in and of itself but fundamental to improved discoverability. Based on our own experiences and the outcomes of the Network Task Forces on semantic enrichment and data quality we will make efforts to improve our aggregation tools to clean and normalise incoming and existing metadata. We will also develop and deploy a component that crawls the media linked to from the metadata provided to us, extracts technical metadata from the media files and stores and index those metadata for search.

To improve the quality the quality, quantity, richness and multi-linguality of added resources and labels we will re-architecture our semantic enrichment module in order to improve the quality and quantity of enrichments. This will provide the basis of a truly multi-lingual search capability, better relevancy ranking in search and availability of browse entry points and facets.

3.5.3 Customer resource management (CRM)

Our current solution for CRM will be reviewed from the perspective that it needs to be much more integrated with our Aggregation tools and APIs.

3.6 Core products

3.6.1 Storage

We will develop storage solutions for user created annotations, clearly separating annotations from originally provided metadata. In order to better support richly inter-related metadata (e.g. hierarchical objects) we will also prototype and implement a storage solution specialised in serving such data with reliability and performance under load.

As part of Europeana 1914-1918 and Europeana Sounds we will develop a SKOS-resource storage and management service. While developed for Europeana it will be developed in such a way as to be possible to be made available as a platform software as a service (SaaS) solution for our Network partners.

Further we will liaise with Europeana Creative partner Kennisland in the developing a service for caching and tiling of high-resolution images. Cached images will be openly licensed or Public Domain. Using this service partners and 3rd-parties will be able to develop reliable and high-performing media rich consumer applications.

3.6.2 Search index

The search index will be upgraded with a new and improved relevancy ranking that will continually be improved by analysing our usage logs. Further we will invest in developing its capabilities to be more quickly updated with user annotations and to support continual publication of ingested metadata.

The index will also be extended in capability and in richness by adding large amounts of multi-lingual labels and contextual metadata resulting from semantic enrichments and dereferencing of SKOSified resources provided by our partners.

3.6.3 Future storage and data sharing (Europeana Cloud)

Europeana Cloud is a major infrastructure project led by Europeana. Its goal is to provide a new platform for metadata and content aggregation for Europeana and other European aggregators. The core ability of the platform will be storing format-agnostic metadata and content records on behalf of aggregators and providers and allowing them to build customized aggregation flows based on it.

The envisioned economies of scale will allow aggregators to gain efficiency through cutting costs on IT and development efforts and outsourcing a major part of these efforts to the body governing the shared infrastructure.

Europeana is going to lead by example and adopt the infrastructure as soon as it matures enough to substitute the proprietary storage of the Europeana today. The main part of this move will occur in 2015 but already in 2014 we will start experimenting with the new storage infrastructure and possibly migrate parts of our current storage systems to it.

3.6.4 Hosting & Systems

Efficient and powerful hosting is a necessary pre-condition for stable operation of the entire Europeana digital infrastructure. In 2014 Europeana will choose a new hosting provider due to the expiration of the contract with the current main hosting provider. We will use this opportunity to fundamentally upgrade our vision about the role of hosting and, consequently, create a new set of requirements for choosing a new hosting provider.

The main themes of the new set of requirements will be native support for continuous integration by the system, natural connection between all the environments - development, testing, acceptance and production - within a unified technological platform and cloud-based

D5.3: Product Development Plan

infrastructure. With the new hosting we aim at better utilizing our development resources, providing the Europeana Network with better abilities to participate in the building of the platform and gain more transparency into the hosting costs.

4 Product Roadmap

	Product	Q2 2014	Q3 2014	Q4 2014	Q1 2015	Q2 2015
B2C	Portal, Search UI & Landing Page	Multi-lingual search & display (Eu2) Hierarchical objects search v1 (Eu2)	Web Style Guide implementation (Eu3)	Agents and Subjects browse & search (Eu3)		
	Portal, Object Pages	Hierarchical objects display v1 (Eu2) Contextual information display (Eu2) Embedded TEL Newspaper viewer (Eu2) Support for standard media embeds (Eu2)	Web Style Guide implementation (Eu3)	Agents and Subjects display (Eu3 T2.1.4)	Improved WebResource display (Eu3)	Public tags and image annotations (based on Creative T2.2.4)
	Portal, My Europeana	Multilingual preferences/settings (Eu2)				Personal annotations management Annotations moderation for admins
	Exhibitions	Research as part of Channels (Sounds WP4)	Research as part of Channels (Sounds WP4) Omeka as Channels CMS report (Sounds WP4)	Exhibitions strategy and development plan		
	End-user blog		Web Style Guide implementation			
	1914-1918 (sub-contract)	v3.1 Release (end of May)	Web Style Guide implementation Improved contribution forms RunCoCo as Channels CMS report (Sounds WP4)	Batch upload of media to Internet Archive for digital preservation Improved media viewers RunCoCo framework update		
	1989 (Awareness)		89 Voices Web Style Guide implementation			

D5.3: Product Development Plan

	Open Culture (sub-contract)	v2.1 Release (mobile-friendly)				
	Channels (Sounds)	Request for proposal/ contracting designers UX research & design	UX research & design Technical design First (internal) prototype	Second (internal) prototype	First working prototype - Music and Map themes	Channels extensions
B2N	Europeana Pro (sub-contract)	UX research & design	UX research & design Development New Pro Beta	DMS integration Full migration and 2.0 release including DMS (Eu3) Integration of ENUMERATE		
	Newsletters	Web Style Guide implementation Release of updated newsletters incl. first Network Newsletter	Developer Newsletter			
	CRM		CRM Review			
	Statistics Dashboard (sub-contract)	Prototype (Eu2, sub- contract Journalism+ +)		Version 1 (Eu3, sub- contract) - featuring increased automisation, partner specific dashboards, and ENUMERATE stats		
B2D	APIs & Linked Data (Creative)	Labs integration, migration & documentation (Creative) Provider & Dataset API (Inside)	Content re-use framework enabled API prototype (Creative, T2.1, T2.2.3) Semantic Enrichment Service (Eu3, CNR- ISTI) Image Similarity Service API (Creative, T2.2.1, MS16, AIT) OAI Server (Creative T2.1)	Content re-use framework enabled API (Creative, T2.1, T2.2.3) Image embed service (Creative, KL) Synched Triple Store (Creative T2.1)	Metadata Selection & Export Service (Creative) SKOS repository service (Sounds T5.1)	Open Source Triple Store and SPARQL endpoint
	Europeana Labs (Creative T1.2)	Research & Design of version 1.0 Development and release of version 1.0	Research & Design for version 2.0 Developing version 2.0 (for a Q1 2015 release)	Researching and developing version 2.0 (for a Q1 2015 release)		
	GLAM-Wiki Toolset	Release v1	Release v1.1			

D5.3: Product Development Plan

	(Wikimedia Chapters)					
LOB	UIM	<p>Continuous publication (PSNC)</p> <p>Hierarchical objects (Eu2)</p>	<p>Common Dataset and Metadata repository (NTUA)</p> <p>Interface between Europeana and Content Providers (NTUA)</p> <p>Content re-use framework integration into the UIM (Creative, T2.1)</p>	Global Dataset Reports (NTUA)	User annotations batch import (Eu3 - Historypin, Sounds T5.5)	Roundtripping of data between Historypin and Europeana
	Enrichment & Data quality	<p>Improved Semantic Enrichment (Eu2, CNR-ISTI)</p> <p>Improved validation rules for EDM (Eu2)</p>	<p>Improved Semantic Enrichment (Eu3, CNR-ISTI)</p> <p>Improved validation rules for EDM (Eu2)</p>	Automatic quality classification of metadata (Eu3)	Data validation, cleaning & normalisation (Sounds T5.4)	
Core	Storage	<p>Hierarchical Objects (Eu2)</p> <p>User annotations - object tags and image annotations (Creative T2.2.4 MS22, AIT)</p>	<p>Image Caching and Embedding Service (Creative, KL)</p> <p>Contextual class instances storage (Eu3, CNR-ISTI)</p> <p>User annotations - sets (Awareness - DSP)</p>	SKOS repository (Awareness, Sounds)	Support for EDM-extensions (Sounds, Fashion)	
	Index & Search	<p>Hierarchical Objects (Eu2)</p> <p>Improved relevancy ranking based on log analyses (Eu2, CNR-ISTI)</p>	<p>Analysis and prototyping of improvement suggestions from the WebWeaving and UVA-reports (Eu3, CNR-ISTI)</p> <p>Search recommendations (Eu2, CNR-ISTI)</p>	<p>Contextual class instances search (Eu3, CNR-ISTI)</p> <p>Rapid re-indexing framework (Eu3, PSNC)</p>	Support for EDM-extensions (Sounds, Fashion)	
	Hosting & Systems	Migration to PaaS preparation (Eu2)	Migration to PaaS (Eu3)			
	Future storage (Cloud)		Prototype Service v.2	Prototype Service v.3 (first production use in Europeana)		