



DELIVERABLE

Project Acronym: Europeana Cloud
Grant Agreement number: 325091
Project Title: Europeana Cloud: Unlocking Europe's Research via The Cloud

D7.6 Periodic progress report (Final Year, Month 30, July 2015)

Revision: FINAL

Authors:

Alastair Dunning, Europeana
Els Jacobs, Els Jacobs, Els Jacobs Advies & Onderzoek

Project co-funded by the European Commission within the ICT Policy Support Programme		
Dissemination Level		
P	Public	

Revision History

Revision	Date	Author	Organisation	Description
1.0	30/7/2015	EJ, NE, AD	Europeana, EJ Onderzoek	

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This deliverable contains original unpublished work except where clearly indicated otherwise.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.

Period 3 Description of Work and Results so Far

Project Objectives

- 1. To provide access, at Europeana, to 2.4m new metadata records and 5m research focussed items from across European Universities, libraries, data centres and publishers**

Via the aggregator in the project, The European Library (TEL), 2.31m records have now been aggregated from partners in the project. 1.54m of these are present in Europeana. The lack of clear rights statements in data from the contributing partners (particularly, the aggregator Bielefeld Academic Search Engine (BASE) means that the project is still deciding on how best to make this metadata available via Europeana.

Work on the content ingestion plan is now taking place. Since there have been significant staff changes at The European Library, new ingestion routes have been decided upon for this data and the timetable for ingestion extended. For the newspapers cited within the Deliverable 4.2 Content Ingestion Plan, they will follow the ingestion route via The European Library, while for other formats Work Package 4 partners will upload content directly to Europeana Cloud (using the tutorial developed – see below).

- 2. To create a cloud based infrastructure capable of delivering cost-efficient content and metadata storage for stakeholders across Europe**

The technical architecture described in Deliverable 2.2 Europeana Cloud Architectural Design has been the guide for the ongoing development of Europeana Cloud (as led by Work Package 2). Within this period, the full software (as developed so far) has been made available on GitHub (<https://github.com/europeana/Europeana-Cloud>). This is accompanied by a tutorial explaining how to upload data to Europeana Cloud API, <https://github.com/europeana/Europeana-Cloud/wiki/Europeana-Cloud-API>.

More of the services that acts as building blocks underlying the file sharing and storage system of eCloud have now been completed. These include the Data Processing and Verification Service, the Data Look Up Service and the Asynchronous Messaging Services.

Four of these building blocks are already in place – the Metadata and Content service that allows users to upload data; the Unique Identifier Service that gives each piece of data a unique identifier; the Authentication and Authorization Service; and the Logging Service.

There has also been planning on the new services that will be added to allow specific operations to be undertaken by customers using Europeana Cloud. These mostly allow customers to perform particular actions over EDM (Europeana Data Model) metadata. These are an EDM Search Service, EDM Enrichment Service, EDM Validation Service, EDM Normalisation and Cleaning Service. There will be also a Content Analysis Service (to

check the quality of digitised content, eg the resolution of a digital images) and an International Image Interoperability Framework (IIIF) compliant Image Server to allow for publication of high resolution images. These services will be rolled out between January and April 2016.

3. To understand and incorporate the legal, strategic and economic issues of a cloud-based system for content for cultural heritage institutions and domain aggregators

Following on from the publication of the Deliverable 5.6 Business Model and the Deliverable 5.7 Product Requirements at the end of Year 2, work has focussed on continuing the relationship between Europeana Cloud and its potential aggregator customers. Further surveys have been issued to gain more understanding of the aggregators' infrastructures, particularly the costs of their existing hardware.

A first version of Deliverable 5.2 Handbook for the Europeana Cloud participants has been completed. Renamed the Europeana Cloud User Guide, it provides an outline of how customers can work with Europeana Cloud, and what services it offers. In the final part of the project more work needs to be done to connect the User Guide to the services mentioned above, and also to the existing technical document in GitHub.

Deliverable 5.3 Revised Europeana Licensing Framework has also been completed. Following discussion with partners in the project and potential customers outside the project, this lays out the access framework for Europeana Cloud, ie how data uploaded by customers can be annotated and used by other customers of the infrastructure, and indeed by third parties.

There has been ongoing work on consideration of the Governance Structure for Europeana Cloud. Various options have been discussed and proposed (including at the eCloud Plenary in Edinburgh). A cooperative association is one potential solution - a Consortium Task Force to discuss further will present its recommendations by October.

4. To achieve a broad consensus among European content aggregators and research networks on the advantages of a cloud based solution

Work during Year 2 resulted in an additional milestone that summarised aggregators needs and fed into the Business Model and the Product Requirements (D5.6 and D5.7). In the past six months, this consensus has been used to refine the technical product and relevant documentation, and the ongoing evolution of the business model.

5. To develop a digital platform, named Europeana Research, to discover and use European research

Perhaps the most significant achievement of the past six months has been the launch of the Europeana Research platform (<http://research.europeana.eu>). This has helped begin the process of forming a community of researchers around Europeana content. Featured

collections have been added to indicate to the research community which Europeana data could be valuable for them. Such work has also provided the framework for more systematic inclusion and interpretation of Europeana data for research use. This will happen in the final six months of the project.

6. Via this cloud to provide tools and services for researchers that permit innovative research that exploits digitised content in Europeana

WP3 has begun work on the third set of tools, exploiting Europeana's collection of digitised historic newspapers. Over 40 research groups responded to a call to provide research questions to exploit the newspapers; around half of them attended a workshop in Amsterdam. This has now provided the background data to create a tool to exploit the newspapers; initial prototypes have focussed on creating a tabletop interface to better analyse the newspapers corpus.

Recent Events and Presentations

Events

Event Name	Event Description	Location	Dates	Number of Participants	Website
"So We've Built It, But Have They Come?" An eCloud/NeDiMAH API Workshhop	Workshop organised by TCD for WP1 work in APIs, bringing together expert DH researchers, 'keen-but-green' researchers and developers to discuss API use in research within the Cultural Heritatre sector	KB, The Hague, Netherlands	17th December 2014	16	n/a
WP2 Dev Meeting	<p>Overall product vision of cloud along with some timeline of the year - to communicate a shift from a project mindset to a product one.</p> <p>Going through each of existing services and discussing its current development state, documentation etc., listing what is missing in thing which were finished, listing what should be added.</p> <p>Result: updated information in JIRA - about existing services.</p> <p>Going to production - discussing what is needed to start delivering ECloud in production mode, incl. software releases plan, hosting environment, possible SLA, upgrade procedure for production environment, public documentation etc. and also migration of each interested partner to ECloud (both scope and timing).</p> <p>Result: updated information in JIRA - about tasks necessary to achieve production-ready state of the service.</p> <p>Defining roadmap for the last year of the project.</p> <p>Result: sprints list with draft goal of each sprint until the end of the project + detailed plan of the next sprint.</p> <p>Discussing more detailed plans for development of new services, including choosing technologies (backlog grooming).</p> <p>Result: updated information in JIRA - about developments planned until the end of the project.</p>	Poznań, Poland	9-11 March 2015	12	
Europeana Research Workshop on Tools, Services and Content Priorities in Archaeology and the Classics	Workshop on tools user requirements and content recommendations for Archaeologists.	London, UK	23-7-2015	20	

Presentations

Event Name	Event Description	Location	Dates	Number of Participants	Website
CERL Coordinating Committee	meeting, presented progress report	ÖNB, Vienna	16-17 March 2015	20	www.cerl.org
CERL Board of Directors	meeting, presented progress report	Bodleian Library, Oxford	16-4-2015	15	www.cerl.org
IDCC15	10th Anniversary conference for Digital Curation Centre (TCD Presenting)	30 Euston Square, London	9-12 Feb 2015	100+	http://www.dcc.ac.uk/events/idcc15
Blue Skies Above, Solid Ground Below: Innovation and Sustainability in Digital Humanities	1st Digital Humanities early career conference (King's College London) - TCD, UGOT and DCU poster presentation of work conducted in the context of WP1 on Case Studies	King's College London, Strand Campus, London	18-6-2015		https://blueskiessolidground.wordpress.com/