



Europeana – Core Service Platform

DELIVERABLE

D1.1: WORK AND IMPLEMENTATION PLAN TO INNOVATE THE AGGREGATION INFRASTRUCTURE

Revision	9
Date of submission	29 January 2016
Author(s)	Henning Scholz (EF), Cécile Devarenne (EF)
Dissemination Level	Public



Co-financed by the European Union
Connecting Europe Facility

REVISION HISTORY AND STATEMENT OF ORIGINALITY

Revision History

Revision No.	Date	Author	Organisation	Description
1	01/12/2015	Henning Scholz	EF	ToC
2	18/12/2015	Henning Scholz	EF	Introduction and expert hub chapter drafted
3	08/01/2016	Joris Pekel, Cecile Devarenne; Nienke van Schaverbeke, Adina Ciocoiu	EF; TEL	Two more chapters added; expert hub chapter reviewed and updated
4	15/01/2016	Henning Scholz, Cecile Devarenne; Valentine Charles	EF; EF/TEL	Review & amendments, completing chapters, alignment of chapters, adding roadmap
5	17/01/2016	David Haskiya, Pablo Uceda Gomez, Jeroen Cichy	EF	Review of the document
6	21/01/2016	Henning Scholz, Cecile Devarenne	EF	Update and amendment to incorporate reviewers comments
7	26/01/2016	Marzia Piccininno, Maria Teresa Natale; Kerstin Herlt; Gisela Baumann; Kate Fernie; Jef Malliet; Stephan Bartholmei; * Jill Cousins, Harry Verwayen, David Haskiya, Victor-Jan Voss, Antoine Isaac, Henning Scholz, Cecile Devarenne	MICHAEL; ACE; FUB; 2Culture; Erfgoedplus.be; DDB; EF	External review by aggregators, internal review by EF, amendments and rephrasing to incorporate reviewer suggestions, consistency check
8	28/01/2016	Beth Daley	EF	Language review
9	28/01/2016	Henning Scholz	EF	Final version

* All Europeana DSI domain and thematic aggregators plus two national aggregators have been invited to review the document. Listed here are those partners that made contributions and suggestions for changes.

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This deliverable contains original unpublished work except where clearly indicated otherwise.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. The sole responsibility of this publication lies with the author. The European Union is not responsible for any use that may be made of the information contained therein.

Table of Contents

1 Purpose of this document	4
2 Summary	4
3 Towards a new aggregation infrastructure.....	5
3.1 Content strategy.....	8
3.2 Expert hubs.....	8
3.3 Shared technical infrastructure in the mid-term	8
3.4 Towards a radically different technological approach in the long-term.....	10
4 From aggregators to expert hubs	10
4.1 Developing the expert hub concept.....	11
4.2 The expert hub role.....	11
4.3 Transforming aggregators into expert hubs.....	12
4.4 Criteria, key activities and key values for expert hubs	12
4.5 A role for national aggregators	15
5 A shared technical infrastructure for the Europeana DSI.....	16
5.1 Towards a collaborative set of requirements: analyses of the existing aggregation tools and workflows.....	16
5.1.1 Analysis of the current situation	16
5.1.2 Aggregator consultation.....	17
5.1.3 Europeana Foundation and The European Library (TEL) use cases.....	21
5.2 Towards a new set of tools dedicated to metadata experts: Metis for The European Library and the Europeana Foundation.....	24
5.2.1 Business requirements phase	24
5.2.2 Development of Metis	26
5.2.3 Technical design principles of Metis	27
5.3 Streamlining data quality related processes between the Europeana Foundation and Europeana DSI aggregators	29
5.4 The future of Metis.....	30
5.4.1 Metis, a set of data ingestion tools to be used by all expert hubs.....	30
5.4.2 Metis, a set of data services connected to external systems.....	31
5.5 Beyond Metis	31
6 A content strategy for the Europeana DSI	32
6.1 Development of the Europeana database over time.....	32
6.2 Towards a content strategy	33
7 Roadmap and conclusion	34
Appendix.....	37
Definitions.....	37

1 Purpose of this document

This document describes the key elements of a planned new aggregation infrastructure and a roadmap to develop and implement it over the coming years. It incorporates the outcomes of several workshops and virtual meetings as well as consultations and one-on-one sessions with the aggregating partners of the Europeana Digital Service Infrastructure (DSI).

2 Summary

The Europeana DSI is using an aggregation infrastructure that was developed more than five years ago. While a few key elements changed along the way (data model, metadata licensing, ingestion tools), the principles of the aggregation infrastructure are still the same. This includes the way the current aggregator model supports a quantitative growth of the database instead of improvements of data quality. However, technology and demands have changed and so must the way data is collected and shared by cultural heritage organizations. This does not only affect the technical infrastructure but also the way we collaborate on organizational and individual levels.

We need to transform the way we make Europe's cultural heritage available by turning the aggregator model upside down: from a hierarchically organized top-down approach, we need to change and start collaborating as interconnected nodes that support each other and work together to provide value to Europe's memory institutions. The expert hub concept is key to this change: domain and thematic aggregators will become expert hubs of the Europeana DSI, recognizing the expertise they already provide and allowing for an increased emphasis on expertise-based services. The technical infrastructure should facilitate this change and allow us to bring the highest possible quality of Europe's rich cultural heritage online. Together with the national, thematic and domain aggregators, we aim to create shared, dynamic, efficient and cost-effective metadata aggregation for the Europeana DSI.

As part of this innovation in the sector and in consultation with the Europeana DSI aggregators, we are developing Metis, a toolset for harvesting, analysing, transforming, enriching and publishing data on the Europeana platforms (e.g. Europeana Collections). The Europeana Foundation and The Europeana Library are the first two customers for this development. The back-end services developed for Metis in 2016 will be available for other partners to use. At a later stage (2017 onwards), Metis will be developed further and offered as a package of tools to more expert hubs.

In order to prepare for the development of Metis and the expert hub concept, further effort is needed to identify, collate and prioritise the business requirements of the aggregators. Multiple stakeholder consultations with aggregators are scheduled for 2016 and 2017. In parallel, the Europeana Foundation and domain/thematic aggregators will also work with the national aggregators to develop a joint plan on how best to improve and manage the exchange of data between national aggregators, aggregators and the Europeana Foundation as part of the Europeana DSI.

With Metis being used by a number of expert hubs and the generic back-end services being used by other partners, including national aggregators, the responsibility for data publication to Europeana platforms becomes much more shared. The goal for Metis is that data officers working for the expert hubs will be able to publish directly to Europeana platforms without an extra layer of Europeana Foundation data officers doing this for them. This alone will improve efficiency and allow data to be published, updated or removed more quickly.

By 2020, we expect more innovative or revolutionary scenarios or technologies to allow memory institutions that prepare data in a standardized way, following the *Europeana Publishing Framework and Guide*, to publish directly to Europeana Collections. At this stage, expert hubs can focus even more on helping memory institutions, using their expertise and advice to unlock Europe's digital cultural heritage and surface the highest possible quality of digital cultural heritage.

3 Towards a new aggregation infrastructure

Europeana is the network for the cultural heritage sector in Europe, and shares a vision of the world in which every citizen has access to all cultural heritage.¹ To realize this, the Europeana Foundation aggregates a comprehensive, reliable and authoritative collection of Europe's cultural and scientific heritage. Since 2008, the Europeana ecosystem has depended on a network of aggregators who bring together, manage and provide access to data about Europe's cultural heritage. The Europeana ecosystem includes three main types of aggregators:

- National aggregators, whose scope is defined by a specific country or region and whose contributors are situated within that country or region (e.g. German Digital Library).
- Domain aggregators, whose scope is defined by a particular industry sector (such as museums, archives or libraries) and whose contributors are located in more than one country (e.g. The European Library).
- Thematic aggregators, whose scope is defined by a particular topic or theme (such as fashion or food and drink) and whose contributors are located in more than one country (e.g. Europeana Fashion).

Collaborating with aggregators to build a digital collection of Europe's cultural and scientific heritage makes Europeana Foundation an 'aggregator of aggregators'. This means it has been able to scale up quickly, relying on the aggregators as conduits in attracting museums, libraries, archives and audio-visual archives to share their collections on the web (Fig. 1). As a model, this now works with fewer than 150 direct providers contributing material from more than 3,500 memory institutions to the Europeana DSI. But the aggregation landscape has become more and more complex over time and technology has changed enormously since the launch of Europeana in 2008. Audiences now demand more of us: bigger and more beautiful images, playable videos or sound recordings and searchable full text that can be read on any device.

¹ <http://strategy2020.europeana.eu/>

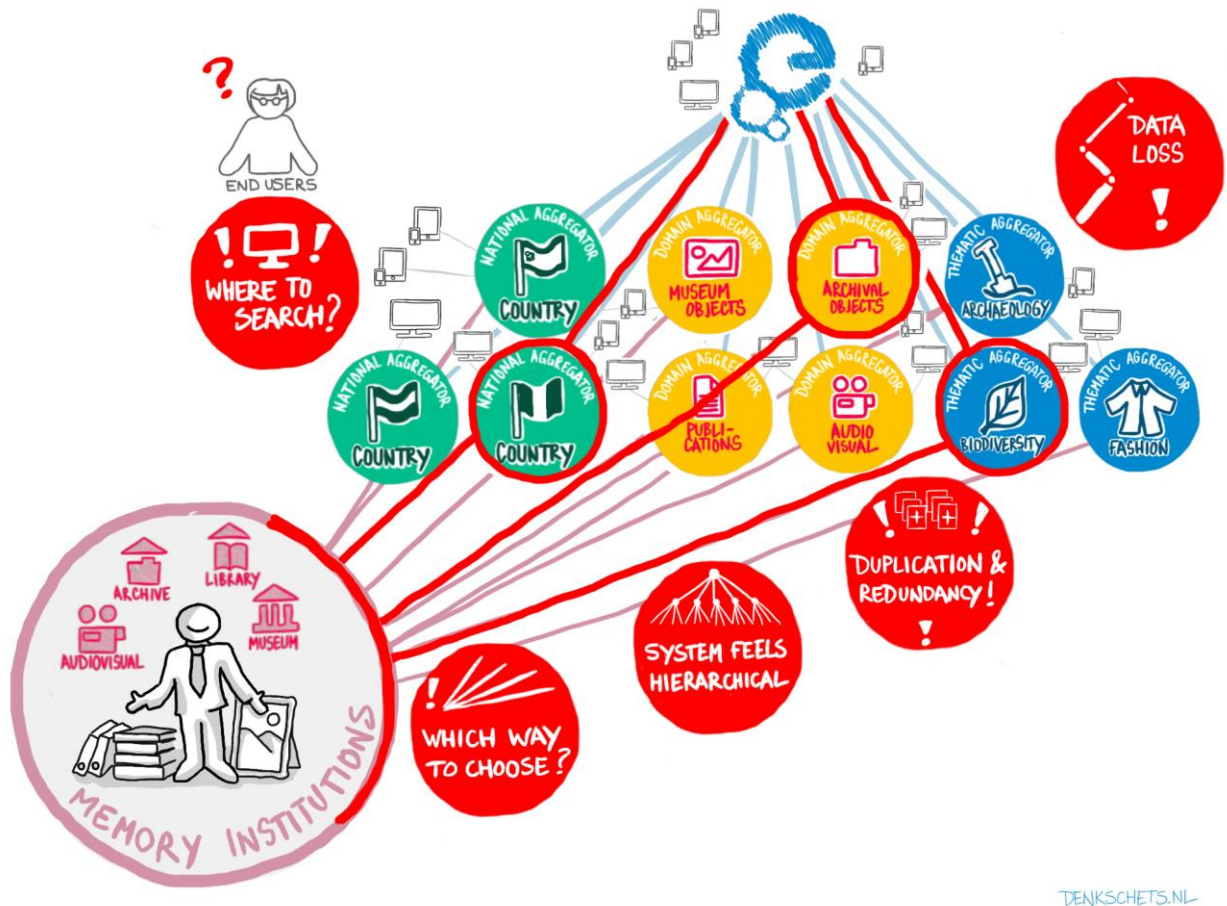


Figure 1. A simplified view of the current aggregator model with a summary of its main weaknesses. The model is simplified in the sense that it only shows one layer of aggregation between memory institutions and Europeana Collections, while in reality there can be multiple layers.

Europeana DSI is now experiencing the drawbacks of the existing aggregator model, such as:

- The current model leads to a long aggregation chain from source to Europeana Collections, which makes it **very slow** to update data. In recent years, it has taken a minimum of one month from source to publication. With continuous publication (see below), this improved to an average turnaround time of two weeks for most datasets.
- Metadata is mapped several times along the aggregation chain which sometimes causes **data loss** and/or loss of precision.
- Memory institutions can choose between several different routes to get their data published in Europeana Collections. This in itself is not an easy decision, but it also causes **duplication** - the same data can be submitted several times and in different qualities.
- The way the aggregator model is conceived and the tools used to process the data require lots of **manual action** to process, publish and update collections.

D1.1: WORK AND IMPLEMENTATION PLAN TO INNOVATE THE AGGREGATION INFRASTRUCTURE

- Some **aggregators cease to exist** after some time, for example when their project funding has ended. This effectively cuts off memory institutions from the Europeana DSI and a new aggregation route needs to be found to allow them to update their data.
- The current aggregator model supports a **quantitative growth** of the database more than improvements in data quality. Data quality overall is still too low and the corpus as a whole is highly heterogeneous, while quantitative growth of the database has continuously increased over the last years. With the current model, it is very difficult to update and improve existing data while new data come in at high pace.
- From an organizational perspective, the current aggregator model is and feels **very hierarchical** and top-down. This does not support a sense of partnership on equal terms between the Europeana Foundation and all aggregators and data partners.

Many of the drawbacks listed above were also reported in the most recent aggregator survey of January 2015.²

Improvements to the current tools and workflows to address some of the issues are ongoing. One important achievement was a major change of the technical infrastructure in summer 2015 that enables the Europeana Foundation to publish data continuously via the Europeana Search API and Europeana Collections. This allows the Europeana Foundation to see the outcomes of ingestion work immediately and to respond to deadlines and the specific needs of data partners more quickly than before.

However, such incremental improvements are not sufficient and there is a need to respond to the needs of the sector with a new aggregation infrastructure. The primary goal for the development of this new infrastructure is the support for **improved data quality**. Several objectives can be drawn up as a direct response to some of the above drawbacks:

- Reduced data loss or loss of nuance between source and aggregation.
- Manual steps a data officer must take must be rationalized away (automation) or reduced (interaction improvements) to free up time for quality improvements.
- Improved statistics reporting (aligned with the *Europeana Publishing Framework* and the planned *Metadata Quality Framework*) and transparency.
- Improved average metadata quality across the corpus.
- Increased use of semantically meaningful and multilingual controlled vocabularies and authorities.
- Data partners enabled to publish directly in Europeana Collections without the Europeana Foundation needing to interrupt beyond a final quality check.
- Data officer collaboration must be supported/improved, e.g. by supporting multiple data officers to work concurrently, whether on parallel or the same datasets.

2

http://pro.europeana.eu/files/Europeana_Professional/Projects/Project_list/Europeana_Version3/Deliverables/EV3%20D1_1%20Aggregation%20Infrastructure.pdf

Beyond the more technical objectives, a new infrastructure should also enable the data experts in our sector to focus more on what they are good at rather than duplicating the work of others. In the process of building a new infrastructure, we must question the principles of aggregation developed in recent years and explore radically new concepts and technologies to connect digital repositories of cultural heritage for every citizen to access.

Building on the recommendations to improve the aggregation infrastructure from March 2015,³ we've identified four components to innovate the aggregation infrastructure:

3.1 Content strategy

The first component is creating a *Europeana Content Strategy*. The strategy will heavily inform the organizational and technological design of the new infrastructure, and allow us to achieve what we need in terms of content development to meet audience needs. This is described in more detail in the chapter [content strategy for the Europeana DSI](#).

3.2 Expert hubs

The second component of the new aggregation infrastructure is to enable a different and better way of collaborating. The current aggregator model is very hierarchical and top-down with Europeana Foundation at the top (see Fig. 1) and everyone else below. We need to change this. The Europeana DSI consists of the Europeana Foundation with its main office in The Hague and a number of aggregators acting as providers of data to Europeana DSI. This includes domain and thematic aggregators, but also national aggregators or platforms. Each of the aggregators and the Europeana Foundation should be considered as interconnected nodes of different sizes and capabilities that support each other and work together to provide value to Europe's memory institutions and citizens: a web, not a pyramid (Fig. 2).

In this context, the expert hub concept can play a key role and is further described in the chapter below ([from aggregators to expert hubs](#)). In recognizing the expertise that is available among the aggregators and that can be shared within the Europeana DSI, we should always be able to find the most suitable support for memory institutions so that they can make the most out of their data.

3.3 Shared technical infrastructure in the mid-term

The third component of the new aggregation infrastructure is a shared set of tools and services to ingest, transform, enrich, manage and publish data (see [a shared technical infrastructure for the Europeana DSI](#)). Maintaining and upgrading a wide diversity of tools requires a lot of (too much) money and human resources. We need to look for commonalities in our aggregation workflows and develop or adapt technical solutions we can use together in order to support data quality improvements while retaining, and ideally lowering, the costs of technical operations. This does not mean that we will develop everything from scratch or deprecate all existing systems.

3

http://pro.europeana.eu/files/Europeana_Professional/Projects/Project_list/Europeana_Version3/Deliverables/EV3%20D1_1%20Aggregation%20Infrastructure.pdf

D1.1: WORK AND IMPLEMENTATION PLAN TO INNOVATE THE AGGREGATION INFRASTRUCTURE

The new tools and services build upon the knowledge gained by the Europeana Foundation, the aggregators and the EuropeanaTech community in recent years. Existing tools and services which have proven to be stable, useful and efficient will ideally be integrated into the new infrastructure, or at least, be enabled to connect to the new tools and services. Apart from reducing the costs of development, adapting existing tools and services should also reduce the costs of retraining and supporting data officers.

The development of the technical infrastructure will be an iterative and incremental process. Aggregators and technical partners will be involved as much as possible in stakeholder consultations, to gather the business requirements and evaluate the services, and to plan for the development of prototypes of other tools connecting to these services.

The tools and services developed for the shared technical infrastructure are planned to enter production from the second half of 2017.

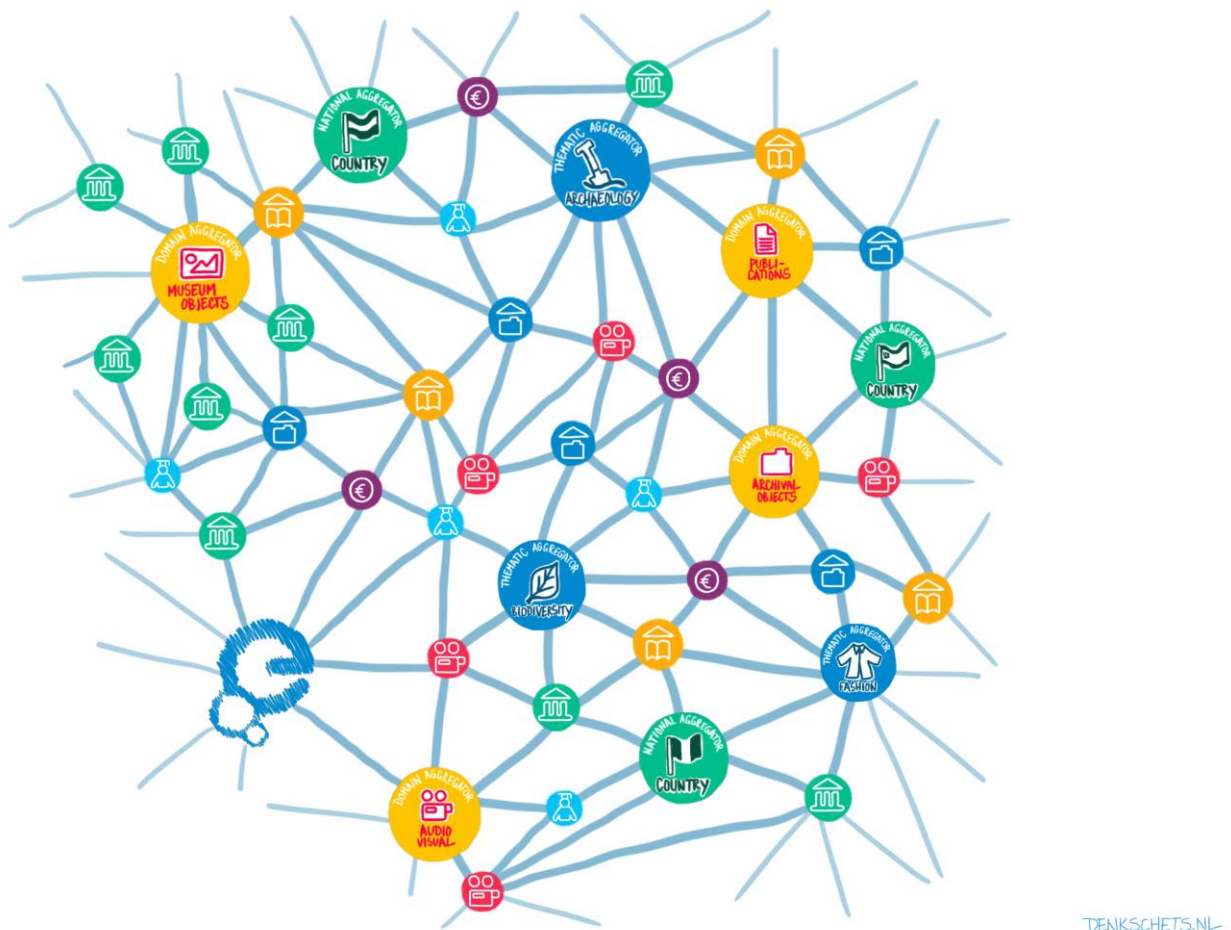


Figure 2. The hierarchical aggregator model should turn into a web of interconnected nodes of different sizes and capabilities that complement each other and work together to provide value to Europe's memory institutions and citizens.

3.4 Towards a radically different technological approach in the long-term

The fourth component of the new aggregation infrastructure is a more visionary approach, looking beyond the current principles and technologies of aggregation or even the new model and technologies outlined in this document. It tries to answer the question of how every citizen can get access to digital cultural heritage and how the Europeana DSI could help if we were to begin developing the systems for that vision today with no constraining legacies, whether organizational or technological. This exploration of alternative data acquisition mechanisms is intended to inform and inspire the development of the new mid-term generation shared infrastructure outlined above but above all to identify and prototype the generation of solutions that one day may replace them.

It is envisioned that this post-Metis generation of services will allow near frictionless exchange of data between libraries, archives and museums and Europeana DSI. Thus largely rationalizing away the need for data officers employed by expert hubs or by the Europeana Foundation to be actively involved in publishing data on the Europeana platform. In such a scenario, the role of data officers would be quality control and data management of the database as a whole.

In the second half of 2016, Europeana DSI will start to put significant effort into exploring alternative data acquisition mechanisms. Using the input for the development of a shared aggregation infrastructure, a group of innovative technology experts will be convened later in 2016 to create proposals for validation by aggregators and technology partners. As this effort is fundamentally one of exploratory research and development, it is difficult to predict at this point when it will result in operational technical services and products. From 2018 onwards at the earliest would be the current rough estimation.

4 From aggregators to expert hubs

When the concept of an aggregator was born, the focus was primarily on collecting, formatting and managing metadata from memory institutions and providing it to Europeana. The reality of 2016 is that 'aggregators' do a lot more than just aggregate data and act as a data provider to Europeana. Domain and thematic aggregators have the knowledge of their domain, the community, the domain-specific metadata standards, etc. needed to support institutions in their domain to make their data available online. They can help enrich and contextualize the data they are working with and support partners in providing curated high-quality data. National aggregators have a responsibility to coordinate a national infrastructure for digital information for the memory institutions in that country.

Therefore, it is time to rethink the concept of an aggregator in order to acknowledge the services that aggregators provide to their partners. In 2015, we started to investigate how domain and thematic aggregators could shift their focus from pure data collection towards becoming expert hubs for their partners. In 2016, we will also investigate the role of national aggregators in this context. Over time, we hope to replace the concept of aggregators with the concept of hubs (e.g.

expert hubs for domain and thematic aggregators), with the consequence that ‘aggregators’ as defined in the *Europeana Glossary of Terms* will no longer exist.⁴

4.1 Developing the expert hub concept

During the meeting of the Aggregators’ Forum on 19 May 2015, it was agreed to set up a Working Group to investigate and develop the concept of expert hubs and create a shared understanding of what this concept of hubs would mean for each aggregator and for the Europeana Foundation.

The Working Group started by creating an expertise and service inventory to get an overview of what all aggregators are currently doing and what they plan to do in the future as an expert hub. Key activities and key values were extracted from this inventory to agree on a lowest common denominator for all expert hubs. Key criteria were developed in order to make it clear what the hard and soft requirements are for an organization to become an expert hub. This is described in more detail below.

As a next step, the key activities identified will be mapped against the initial expertise and service inventory to make the offering as specific as possible and to identify areas of overlap between the hubs to make collaboration as easy as possible and to address duplication of work. The Expert Hub Working Group will also investigate the relationships aggregators have between themselves and other aggregators, their memory institution partners and the Europeana Foundation. At the time of writing, this part of the work is still in progress, as is work on the alignment of roles between expert hubs and national aggregators.

4.2 The expert hub role

Expert hubs are essential for the smooth running and success of the Europeana platform as a Digital Service Infrastructure (DSI) because as a pan-European cross-domain platform, the Europeana Foundation does not have the in-house expertise to support all kinds of memory institutions efficiently. Therefore, domain experts that understand the needs and challenges of memory institutions are required for the Europeana DSI. It needs experts for different metadata schema and application profiles as well as experts for relevant content. It also needs experts who share a similar background and mindset with colleagues from the memory institutions and who can represent these institutions appropriately. Domain and thematic aggregators have increasingly filled this role of experts for the Europeana Foundation over the last years.

Expert hubs should be able to provide domain and/or thematic specific expertise (e.g. on data quality or copyright) and services (e.g. access to documentation, best practices and standards) to the memory institutions they partner with, but also more general, domain-independent expertise and services (e.g. help-desk infrastructure). The general expertise and services an expert hub offers should complement the expertise of other hubs. By collaborating and sharing expertise and services, together, hubs can meet the wide-ranging collective demands of their partners. If memory institutions need services their regular expert hub cannot provide, the hub needs to make sure to help its partner in connecting with another hub that is able to fulfill the needs. An

⁴ See the [definitions](#) further above or <http://pro.europeana.eu/page/glossary>.

example would be a museum that would like to publish their digitized book collection online and has questions about data models for books. The museum would first approach the museum hub. The museum hub would make the connection with the library hub to make professional support available for that museum.

In relation to the expert hubs the Europeana Foundation retains a coordinating role, will call on expert hubs for expertise and access to services it does not itself have. An example is work on data quality or enrichment for which the Europeana Foundation Research and Development team gets in touch with expert hubs to address the issues appropriately.

4.3 Transforming aggregators into expert hubs

The development of domain and thematic aggregators into expert hubs would start with the recognition and definition of the expertise-based services they offer to memory institutions. This will help the aggregators to position themselves in the market and strengthen their brands. It will help to clarify that data aggregation is not the only task and may not even be the core mission of an expert hub. It will also enable the collaboration of expert hubs on specific expertise-based challenges.

The first expert hubs would be launched in 2017 as centres of expertise depending on their maturity and compliance with the criteria (see below). The expert hub concept itself will also be reviewed along the way and adapted if necessary for the next expert hubs being launched.

For the further development of the expert hub and an increased emphasis on expertise-based services, a shared technical infrastructure should be in place for a specific hub. The Europeana DSI is investing in the development of a shared technical infrastructure. This will consist of a set of tools and services to harvest, analyse, transform, enrich and publish data on the Europeana platform (see chapter on [a shared technical infrastructure for the Europeana DSI](#)) and its services for users and re-users.

With this infrastructure in place, making cultural heritage available online and improving its quality will become more efficient, e.g. by reducing manual interventions or by reducing the number of mapping steps, and supporting and promoting increased data quality. Thus the infrastructure will play a significant role in allowing expert hubs to focus more on their expertise-based services and less on maintaining and developing technical services specific to themselves. As a result they should be able to better support their memory institution partners in sharing high-quality data with the world and via the Europeana platform.

4.4 Criteria, key activities and key values for expert hubs

In developing the expert hub concept, the Expert Hub Working Group agreed on some key criteria that need to be fulfilled in order to become a Europeana Expert Hub.

1. Multi-national professional community with national/regional representatives: the expert hub is not restricted to one country or region; it brings together cultural heritage organizations from several European countries. Expert hubs should endeavour to have national/regional representatives acting as a national contact point.

D1.1: WORK AND IMPLEMENTATION PLAN TO INNOVATE THE AGGREGATION INFRASTRUCTURE

2. Domain-specific expertise and knowledge: expert hubs share expertise and knowledge for a specific cultural heritage domain and bring together institutions from that domain. For the time being, we define the limits of a domain-based expert hub on the UNESCO framework for cultural statistics.⁵
3. Legal entity: ideally, one organization must act as the legal entity representing the expert hub. This will facilitate the representation of the hub in project proposals (e.g. Europeana DSI). It will also help with the sustainability of the expert hub.

In addition, the Expert Hub Working Group suggested some common principles for expert hubs:

1. Inclusiveness: offer expertise to both small and large memory institutions; offer a range of possible levels of expertise.
2. Internationalism: collaborate with European and international organisations to increase access to cultural heritage, connect collections and share expertise and best practices.
3. Community: operate based on Europeana Commons Principles⁶
4. Transparency: open about what they do and how.

The principles need more work to be fleshed out and understood by all stakeholders, which will be done in the first half of 2016. That includes e.g. working out what transparency means in practice and how cross-border collaboration is different from the pan-European professional community forming the expert hub.

The Expert Hub Working Group also suggested key activities that expert hubs are envisaged to do. They are deliberately generic as every expert hub will need to specify a work programme specific to the hub for each key activity. This work programme per expert hub is still pending and is expected to be completed for each expert hub with its launch (see also Fig. 3 for an example for activities of The European Library).

Based on the ranking we agreed to date, these activities are most important for expert hubs:

1. Facilitate the work of the domain data partners and their collaboration/communication.
2. Communicate, coordinate and broker between data partners, Europeana DSI, other hubs, projects and other stakeholders.
3. Share knowledge and best practice with domain partners and help implement it.
4. Motivate partners and help them to improve the quality of their data.
5. Raise competency (training, development) of domain partners in relevant aspects of digital cultural heritage (e.g. digital strategy, IPR, metadata quality, tools).
6. Ensure aggregation of data to Europeana DSI.

Based on the ranking we agreed to date, these activities are medium priority for expert hubs:

1. Liaise with domain-specific associations and research infrastructures.
2. Set up, run and facilitate working groups, task forces, forums.

⁵ <http://www.uis.unesco.org/culture/Documents/framework-cultural-statistics-culture-2009-en.pdf> (see page 24 for an overview of domains)

⁶ http://pro.europeana.eu/files/Europeana_Professional/Projects/Project_list/Europeana_Version3/Milestones/Ev3%20MS20%20Cultural%20Commons%20White%20Paper.pdf

D1.1: WORK AND IMPLEMENTATION PLAN TO INNOVATE THE AGGREGATION INFRASTRUCTURE

3. Contribute to editorial boards of Europeana Thematic Collections
4. Represent the interest of the domain (e.g. lobbying support).
5. Maintain, sustain and grow the expert hub.
6. Coordinate the needs and requirements for digital infrastructures within the domain.
7. Support domain partners in getting access to experts in other domains.
8. Support domain partners in developing a digital strategy .
9. Organize and host annual meetings and thematic workshops.

Based on the ranking we agreed to date, this activity is of low priority for expert hubs:

1. Support domain partners exploring data use and re-use scenarios.

The discussion to refine the activities and to finalize the ranking for all potential expert hubs is still ongoing. This also includes ensuring that the wording of the activities is precise and clear.



Figure 3. The European Library envisaged as an expert hub. It displays the library domain-specific activities considered at this time.

Every expert hub should strive to provide value to the memory institutions with which it partners. The Expert Hub Working Group has started to work on value propositions. As the value differentiation is not finished yet for all expert hubs, the below list shows a few examples to indicate the direction this work is taking.

These are examples of proposed value propositions of expert hubs to memory institutions:

1. Improved and semantically enriched data.
2. Increased impact of memory institutions' data.
3. Increased visibility of memory institutions' data and collections.
4. Increased opportunities for content re-use.
5. Increased interaction with memory institutions' cultural heritage online (e.g. re-use).
6. Increased knowledge about project funding opportunities for memory institutions.
7. Increased knowledge about open access and the benefits of data re-use in memory institutions.

4.5 A role for national aggregators

In parallel, national aggregators are searching for common ground amongst themselves. They are exploring possible agreements in a Memorandum of Understanding. There are many common features in their operation, but there are also some differences. They have somewhat different histories and each work within their own policy environments for heritage and digitization.

They can support the Europeana DSI with important services and expertise that complement what more domain-related expert hubs can provide, such as:

- Inclusiveness and balance: to involve memory institutions across all domains, themes and sizes, and to warrant full territorial coverage.
- Synergy: orchestrating or operating the technical aggregation infrastructure and common aggregation services with which the expert hubs may engage.
- Proximity: providing personal support and building strong relationships with data partners and audiences.
- Interoperability: cross-domain expertise.
- De-duplication: help with detecting duplicate objects.
- Sustainability and representation: linking with the member states which determine funding by the EU.
- Language: translation and adaptation of Europeana documents and campaigns to local conditions and language.

More work will be done in 2016 to improve the understanding of common ground among national aggregators and to adjust the complementarity of services and expertise between expert hubs and national aggregators. This should benefit the overall goal of efficiency of the Europeana DSI workflow and avoid duplication of efforts.

5 A shared technical infrastructure for the Europeana DSI

The way we plan to innovate how aggregation is organized informs and interplays with the development of the new tools and technical services for aggregation. Sharing the expertise and experiences gained over the last five years of aggregation must inform how we rethink the way we work and how we will improve the dedicated tools and services we use for data aggregation and management. Analysing what we can use in common and reshuffling the technical infrastructure of aggregating metadata and content in the context of the Europeana Foundation and its network is a challenging task with many moving parts. Focus will be on streamlining the processes and on reducing as much as possible the functional redundancies in the various systems the Europeana Foundation and the aggregators have implemented. We must use our experience to our benefit and adopt more innovative technologies to build a set of common tools suited to the needs of all actors who are part of the aggregation landscape.

The development of the shared technical infrastructure, modelled according to the ‘Three Horizons’ approach to innovation,⁷ consists of:

1. Horizon 1: Maintenance of existing tools and systems like the Unified Ingestion Manager (UIM)⁸ and MINT.⁹ This document does not touch further on these systems as they are in late stages of their product life cycle and have existing well-established procedures for maintenance and any incremental development deemed essential to operations.
2. Horizon 2: Development or adaptation of new and existing tools and services into a shared system of services. Individual tools are interoperable in terms of data exchange. Users of the tools are drawn from multiple aggregators/expert hubs. The development of Metis (see below sub-chapter [5.2](#)) is core to this horizon.
3. Horizon 3: Research, development and prototyping of radically different ways of acquiring data for further distribution (see above sub-chapter [3.4](#)).

As development progresses, successful Horizon 2 and 3 solutions replace the solutions one horizon above, and new horizon 3 research and development projects are initiated.

5.1 Towards a collaborative set of requirements: analyses of the existing aggregation tools and workflows

5.1.1 Analysis of the current situation

The work performed in the first year of the Europeana DSI has enabled us, by analysing existing systems, to better identify needs and begin drafting common requirements between the Europeana Foundation, The European Library (TEL) and other aggregators. Based on an analysis of the results, the following priorities were defined and decisions taken:

- The new ingestion tools must support and promote data quality at all levels of the aggregation process.

⁷ See e.g. <http://blog.hypeinnovation.com/using-the-three-horizons-framework-for-innovation>

⁸ http://pro.europeana.eu/files/Europeana_Professional/Projects/Project_list/Europeana_Version3/Deliverables/Ev3%20D5_2%20Review%20Logical%20and%20Technical%20Architectures.pdf

⁹ <http://mint.image.ece.ntua.gr/redmine/projects/mint/wiki/Wiki>

D1.1: WORK AND IMPLEMENTATION PLAN TO INNOVATE THE AGGREGATION INFRASTRUCTURE

- The new ingestion tools must improve the delivery workflow to Europeana and the data officers' user experience (UX) in order to free up resources to focus on quality.
- The new ingestion tools and infrastructure will build on the results of the Europeana Cloud project and follow a service-oriented modular architecture that allows multiple clients to use the services.
- Development will by necessity stretch over many years and be iterative and incremental. Over this stretch of time, individual tools, modules and services will be released when they are ready (first for testing and then for production).

5.1.2 Aggregator consultation

As stated earlier in this document, recommendations to improve the aggregation infrastructure (published as D1.1 under Europeana Version 3 in March 2015)¹⁰ have highlighted the need to undertake an analysis of the current aggregation landscape, including the various softwares and technical setups used by domain and thematic aggregators of the Europeana DSI to get the data from cultural institutions into Europeana. The technical infrastructure development plan produced in the beginning of the Europeana DSI (published as MS9 in June 2015)¹¹ provides a first inventory of the existing systems in use and the actions for technical improvements planned for the first year of the project. As some tools are already used by several partners, a need was shown to improve the coordination between aggregators in regard to the plans for these tools. For instance, it was recommended that a MINT coordination group would be set up (this group will be formalized in 2016). A workshop was then organized during the Aggregators' Forum meeting in The Hague in May 2015, at which the operations of all aggregators were discussed in order to identify the improvements needed for aggregation tools and workflows.¹²

In order to frame the discussion, categories were suggested to represent various dimensions of the data processing operations:¹³

- Import: ingesting data into a system.
- Transform: transforming data from an input schema to a target schema, edit the data.
- Validate: validating data against a given schema, analyse the data.
- Enrich: appending additional data to ingested data (manually or automatically).
- Publish: giving access to data on a platform (portal, API, etc).

These categories, based on a simple abstraction of the type of operations the Europeana Foundation data officers perform on data, were agreed to be a sufficient and adequate basis to work from and describe the different workflows of all aggregating partners. While operations for aggregation of data to Europeana are similar overall (all or most partners perform harvesting, transformation and enrichment of the data they aggregate), the use cases and approaches differed in their implementation, as described below. With the aim of developing that discussion further and starting to create common requirements, a Working Group has been set up with

¹⁰ http://pro.europeana.eu/files/Europeana_Professional/Projects/Project_list/Europeana_Version3/Deliverables/EV3%20D1_1%20Aggregation%20Infrastructure.pdf

¹¹ http://pro.europeana.eu/files/Europeana_Professional/Projects/Project_list/Europeana_DSI/Milestones/europeana-dsi-ms9-technical-infrastructure-development-plan.pdf

¹² <http://pro.europeana.eu/event/aggregator-forum-spring-2015>

¹³ <http://fr.slideshare.net/Europeana/ingestion-workflows-201505denhaag>

D1.1: WORK AND IMPLEMENTATION PLAN TO INNOVATE THE AGGREGATION INFRASTRUCTURE

aggregation partners of the Europeana DSI only. This Working Group will continue working in 2016 and 2017. In addition to this, to align communication among the partners involved, it is recommended in the draft Europeana DSI-2 proposal that partners share concrete experience on the use of existing tools, by organizing cross-training sessions and residencies between aggregators and the Europeana Foundation.

While all aggregators act as advisers and data experts, they are not necessarily the primary users of the aggregation tools themselves. Several existing scenarios were highlighted in discussions with stakeholders:

- A centralized approach in which memory institutions acting as data providers for Europeana deliver their data to the aggregator and the data officers of the aggregator perform the transformations and enrichments required for delivery to Europeana (and publication of the data in other dedicated platforms).
- A distributed approach with a hands-on mechanism in which memory institutions acting as data providers for Europeana themselves perform most of the data operations on their own collections.
- A custom approach in which cultural institutions perform part of the operations on their data themselves while data officers from the aggregator oversee the process and perform the rest of the operations required for delivery to Europeana.

The table below breaks down each of the above approaches, detailing how they use the data aggregation tools and the strengths and weaknesses of each approach.

Note: Direct providers, national aggregators and other initiatives were added to the table as indicators, they were not part of the consultation in the first year of the Europeana DSI.

	Centralized approach	Distributed approach	Custom approach	Direct provider
User group for the data aggregation tools	Data officers for the aggregator	Any officer from contributing memory institution (data officer, data curator)	Both data officers for the aggregator and any officer from contributing memory institution	Any officer from contributing memory institution (data officer, data curator)
Examples of aggregators/providers	TEL, Europeana, OpenUp!, ¹⁴ Deutsche Digitale Bibliothek (DDB), ¹⁵ Hispana ¹⁶	Europeana Fashion, ¹⁷ CARARE, ¹⁸ Photoconsortium, ¹⁹ MICHAEL, ²⁰ EUscreen ²¹	European Film Gateway (EFG), ²² Daguerreobase ²³	Rijksmuseum, ²⁴ Museo del Prado ²⁵

¹⁴ <http://open-up.eu/en>

¹⁵ <https://www.deutsche-digitale-bibliothek.de/>

¹⁶ <http://hispana.mcu.es/es/estaticos/contenido.cmd?pagina=estaticos/presentacion>

¹⁷ <http://www.europeanafashion.eu/portal/home.html>

¹⁸ <http://www.carare.eu/>

¹⁹ <http://www.photoconsortium.net/>

²⁰ <http://www.michael-culture.eu/>

²¹ <http://www.euscreen.eu/>

D1.1: WORK AND IMPLEMENTATION PLAN TO INNOVATE THE AGGREGATION INFRASTRUCTURE

<p>Role of aggregator (data officers)</p>	<ul style="list-style-type: none"> - Perform all data operations: harvesting; mapping from source to target; enriching; publishing - Advise data providers on what needs to be done - Oversee coordination between the cultural memory institution and Europeana 	<ul style="list-style-type: none"> - Coordinate data operations - Provide user support and act as admins of the tools - Advise data providers on what needs to be done - Oversee coordination between the cultural memory institution and Europeana 	<ul style="list-style-type: none"> - Coordinate data operations - Provide user support and act as admins of the tools - Perform part of data operations (for instance, enriching and publishing) - Advise data providers on what needs to be done - Oversee coordination between the cultural memory institution and Europeana 	<p>Not applicable</p>
<p>Technical support for the tools</p>	<ul style="list-style-type: none"> - Maintenance - Addition of functionalities - Processing of specific tasks in the workflow 	<ul style="list-style-type: none"> - Maintenance - Addition of functionalities - Processing of specific tasks in the workflow - Regular training for users and daily support 	<ul style="list-style-type: none"> - Maintenance - Addition of functionalities - Processing of specific tasks in the workflow - Regular training for users and daily support 	<p>Part of the general support for the collections management system</p>
<p>Examples of tools/platforms</p>	<p>UIM for TEL, UIM for Europeana, BioCASE,²⁶ IPT²⁷</p>	<p>MINT, MORE,²⁸ LoCloud Collections,²⁹ Europeana Connection Kit (ECK)³⁰</p>	<p>D-NET platform³¹</p>	<p>Adlib Museum³²</p>

²² <http://www.europeanfilmgateway.eu/>

²³ <http://www.daguerreobase.org/>

²⁴ <https://www.rijksmuseum.nl/en/explore-the-collection>

²⁵ <https://www.museodelprado.es/>

²⁶ <http://www.biocase.org/>

²⁷ <http://www.gbif.org/ipt>

²⁸ <http://more.dcu.gr/>

²⁹ <https://locloud.pl/>

³⁰ http://www.europeana-inside.eu/documents/eck/eck_documentation.html

³¹ <http://www.d-net.research-infrastructures.eu/node/8>

³² <http://www.adlibsoft.com/products/museum-collection-management-software>

D1.1: WORK AND IMPLEMENTATION PLAN TO INNOVATE THE AGGREGATION INFRASTRUCTURE

<p>Overall strengths</p>	<ul style="list-style-type: none"> - Offer in terms of service: support and operations - Users of the tools are mapping experts and have a global overview of the data they aggregate - Harmonization and integrity of the aggregated data - Homogeneous background of users facilitates the design of the tools 	<ul style="list-style-type: none"> - Any cultural institution can contribute - Data operations are performed by users with perfect knowledge of the source - Knowledge of the target schema prior to use of the tool is not a prerequisite 	<ul style="list-style-type: none"> - Combination of strengths from two previous approaches - Customized solution for a group of users of one domain or country 	<ul style="list-style-type: none"> - Data processed closer to the source: potential for high-quality data - Direct communication between cultural memory institution and Europeana: better understanding of the needs and goals
<p>Overall weaknesses</p>	<ul style="list-style-type: none"> - Requirements on the input data: not all formats or models can be harvested - Extra coordination and communication steps might slow down the process - One actor (data officer) responsible for all steps of the workflow: limited possibilities to process and update collections regularly enough 	<ul style="list-style-type: none"> - Variety of profiles of users: designing tools is a long iterative process - Continuous need for helpdesks in addition to documentation and user support provided by data officers - Risk of lack of full understanding of the target schema or other operations: impact on data quality for the aggregated data 	<ul style="list-style-type: none"> - High level of separate customizations to meet the needs of all domains and countries as well as the various group of users (might result in a lot of manual processes) 	<ul style="list-style-type: none"> - Requires investment and technical resources from the memory institution itself - Amount of data support to be provided by Europeana: this approach cannot be applied to all memory institutions contributing for reasons of scalability

All the approaches reviewed have their strengths and weaknesses. In the case of a distributed approach, any memory institution interested in contributing can be targeted, supporting the Europeana mission to be inclusive and aggregate data from all cultural memory institutions who would wish to contribute. The offered tools are designed to be user-friendly, enabling data operations to be performed directly at the source by users with varying technical data expertise and with data curators being fully active in the process of transforming their data. Yet, considering the variety of profiles of users, designing efficient tools is a resource-intensive iterative process and continuous user support for these tools has a significant cost in terms of resources, which

cannot be reduced over time. For instance, as observed with the MINT tool used within the LoCloud project:

“the transformation and the publication of providers’ metadata to the selected target schema (LIDO, Carare2.0, EDM) is a process during which providers may encounter difficulties, even if they are well aware of MINT’s functionality. This is because the majority of the providers do not have a strong technical background and it may be hard for them to follow the ingestion workflow combining MINT’s functionalities to reach the desired result.”³³

In the case of the more centralized approach, a minimal set of requirements is given to cultural institutions for delivery of source data into the aggregation system (for instance, a list of supported metadata schema), and further operations on data are performed by data officers who are expert in, if applicable, the domain or theme they aggregate for, data processing and the target schema for aggregation. The fact that their backgrounds are homogeneous helps in designing adequate systems. It is desired that the users of these systems can process data with as little technical support as possible.

A shared technical infrastructure, if understood as one set of tools used by all (data providers as well as data aggregation officers) and fulfilling various specific needs, appears to be a major challenge with a high risk of a solution resulting in a ‘one size fits no-one’ outcome. A set of tools dedicated to data officers will be developed first in a process that comprises continual consultations with stakeholders and continual testing by users (data officers).

5.1.3 Europeana Foundation and The European Library (TEL) use cases

The Europeana Foundation and The European Library perform their tasks for aggregation using a common technical framework implemented differently to match both use cases. The ‘Unified Ingestion Manager for The European Library’ and ‘Unified Ingestion Manager for the Europeana Foundation’ were released in 2013 and have been further improved and stabilized since then.³⁴ In 2015, the decision was made to gradually integrate The European Library’s data services into Europeana’s platform. As a first step in this transition, an analysis and comparison between the two systems and workflows was carried out in order to envision how a shared system could be used by both The European Library and the Europeana Foundation, keeping in mind also the longer term need for a shared solution for other Europeana DSI aggregators.

The European Library acts as the library domain aggregator, while the Europeana DSI is an aggregator for already standardized aggregated data, resulting in different scope and needs. Shared functional requirements have been identified but so have diverging requirements with some workflows entirely specific to one or the other. This is likely to be the case for most of the domain aggregators compared to the Europeana Foundation and is a major design and development challenge to overcome.

³³ <http://www.locloud.eu/content/download/1003/6960/version/1/file/LoCloud+D2+2.pdf>

³⁴ http://pro.europeana.eu/files/Europeana_Professional/Projects/Project_list/Europeana_Version3/Deliverables/Ev3%20D3.2%20ProductDevelopmentReport.pdf

D1.1: WORK AND IMPLEMENTATION PLAN TO INNOVATE THE AGGREGATION INFRASTRUCTURE

		The European Library	Europeana Foundation
User profile		Data officers; expertise in library data content and metadata formats, good knowledge of the Europeana Data Model (EDM)	Data officers; expertise in the Europeana Data Model (EDM), generic knowledge of cultural institutions' data
Type of data and content managed within the infrastructure		Most common bibliographic standards (MARC, UNIMARC, METS, MODS) and EDM ³⁵ for delivery to Europeana. Management of full-text in various formats (e.g. METS-ALTO, PDF)	Data in EDM
Dataset and organization management	System	SugarCRM ³⁶ customized for TEL's specific needs	SugarCRM customized for Europeana's specific needs
	Specifics	All history of executed tasks on datasets is maintained, exports of ingestion plans, collections descriptions, full implementation of relations between organization, dataset and project entities, synchronization mechanism	Data Exchange Agreement information maintained, basic inventory of datasets with status and total amount of ingested records filled in manually
Workflow orchestrator	System and specifics	UIM specific implementation, Solr as storage engine, common component (workflow execution mechanism based on OSGi)	UIM specific implementation, Mongo as storage engine, common component (workflow execution mechanism based on OSGi)
Import	System	REPOX ³⁷	REPOX
	Specifics	<ul style="list-style-type: none"> - Harvesting protocols: FTP, HTTP, OAI-PMH - Input format: xml - Input schema: any - Incremental harvest supported 	<ul style="list-style-type: none"> - Harvesting protocols: HTTP, OAI-PMH - Input format: xml - Input schema: only Europeana Semantic Elements (ESE) and EDM are used - Incremental harvest functionality unused.
Transform	System	Automatic transformation based on XSLT mappings. XSLT editor available within UIM for update on XSLT.	MINT, customized implementation for Europeana

³⁵ <http://pro.europeana.eu/share-your-data/data-guidelines/edm-documentation>

³⁶ <https://www.sugarcrm.com/>

³⁷ <http://pro.europeana.eu/blogpost/introducing-repox-a-tool-to-manage-metadata-spaces>

	Specifics	Automated transformation from input to Internal Object Model (IOM), edition of xslt for customization of mapping and changes in the data on a dataset basis; history of xslt editions maintained	Manual mapping from EDM External/ESE to EDM Internal ³⁸ for each dataset created in a user interface by drag and drop; mapping json file and xslt generated; all history of mappings is maintained
Enrich	System	Custom UIM plugin	Custom UIM plugin using the Europeana Semantic Enrichment service ³⁹
	Specifics	Described in the inventory ⁴⁰ produced by the Evaluation and Enrichment Task Force ⁴¹	Described in the inventory produced by the Evaluation and Enrichment Task Force
Publish (data out services)		TEL portal, ⁴² Newspaper data dumps service, OpenSearch API, ⁴³ OAI-PMH server, ⁴⁴ Linked data exports	Europeana APIs ⁴⁵ (including REST APIs, OAI-PMH server, Europeana Linked Open Data), Europeana Collections ⁴⁶

Both versions of the UIM are stable in the sense that data is aggregated daily to be outputted via the various interfaces the Europeana Foundation and The European Library make available.

However, over the years a series of drawbacks has been identified in both versions:

- Processing one collection into either of the versions is a cumbersome process:
 - Each workflow (consisting of one or several operations on data) has to be triggered manually.
 - Whenever changes are made within part of a collection, it needs full reprocessing which makes the operations slower than they need to be.
 - The frequency of updates for a given collection is considered too slow. On the Europeana Foundation side, publication of data used to happen on a monthly basis, which had the consequence that data published after aggregation to Europeana were not necessarily up-to-date. The implementation of a continuous publication mechanism under the first year of Europeana DSI has been considered as a major improvement already but further optimization of the processes is still very much needed.

³⁸ See documentation for EDM Internal schema:
<https://github.com/europeana/corelib/wiki/EDMObjectTemplatesEuropeana>

³⁹ <http://144.76.50.251:8080/enrichment-framework-gui-0.1-SNAPSHOT/>

⁴⁰ http://pro.europeana.eu/files/Europeana_Professional/EuropeanaTech/EuropeanaTech_taskforces/Enrichment_Evaluation//AppendixA_InventoryEnrichmentTools_102015.pdf

⁴¹ <http://pro.europeana.eu/taskforce/evaluation-and-enrichments>

⁴² <http://www.theeuropeanlibrary.org/tel4/>

⁴³ <http://www.theeuropeanlibrary.org/tel4/access/data/opensearch>

⁴⁴ <http://www.theeuropeanlibrary.org/tel4/access/data/oaipmh>

⁴⁵ <http://labs.europeana.eu/api>

⁴⁶ <http://www.europeana.eu/portal/>

D1.1: WORK AND IMPLEMENTATION PLAN TO INNOVATE THE AGGREGATION INFRASTRUCTURE

- In addition to the redundancy of the process itself, data officers perform most of their analysis of the data manually, and lack the necessary information and tools to be able to support the desired increased focus on data quality.
- The consequent of the two previous drawbacks is that the aggregation workflows appear as 'black boxes' to external partners - they have no sight of what is happening to their datasets when they are being processed by the Europeana Foundation. This lack of transparency obviously has a negative impact on communication within the full aggregation landscape.
- The amount of data aggregated has considerably grown over the years and while they could successfully be ingested, technological limitations of the systems have surfaced. Maintenance of the two infrastructures has proven to be costly and can only be performed by the in-house development team. Working in collaboration with technical partners towards a shared infrastructure cannot happen with the system as implemented.

Resulting from this analysis, the choice has been made to develop a new set of tools to replace the UIM which can serve both the Europeana Foundation and The European Library. Development began in January 2016. The first iteration of Metis combines two groups of users with different expertise and work scope, making it a true test of the feasibility of one centralized tool serving two different types of aggregators and use cases.

5.2 Towards a new set of tools dedicated to metadata experts: Metis for The European Library and the Europeana Foundation

5.2.1 Business requirements phase

Following the analysis of the existing systems in use by The European Library and the Europeana Foundation, areas of collaboration were identified where workflows or processes could be shared, and specific aspects related to the differences of scope between The European Library and the Europeana Foundation were investigated. For these, decisions had to be made about whether the functionalities should be ported to the new system or if some of the services in place could be deprecated. A transition plan for The European Library was created as part of their strategy for 2015. The objective as described in the plan is to gradually transition the The European Library data services into a shared infrastructure with the Europeana Foundation. The following will be retained: the tools ensuring the aggregation of data (both for catalogues and digital objects) and the services for distributing this data to the Europeana platform and third parties such as research infrastructures.

As part of the common requirements work, the following areas were identified and discussed:

- **User experience and user design:** It was acknowledged that insufficient effort has been put in to designing the existing tools for the Europeana Foundation and The European Library aggregation. As user experience has a direct and obvious impact on the efficiency of the system, it will be crucial for the development of Metis to perform UX research, firstly with The European Library and the Europeana Foundation data officers, and secondly with other data officers of Europeana DSI aggregators.
- **Dataset and organization management:** For both organizational and archival purposes, information about datasets and organizations needs to be recorded and preserved. Both

The European Library and the Europeana Foundation use their specific customizations of SugarCRM to maintain administrative metadata as well as business processes, but further standardization of the way administrative metadata is kept is a priority. This could be supported by implementing the existing EDM profiles for organization and dataset.⁴⁷ Making part of this information (for instance, the publication status of one dataset) visible to providers is considered as one of the needs to meet. A Europeana provider and dataset API⁴⁸ is in place already, developed in the context of the Europeana Inside project⁴⁹ and designed according to the mentioned EDM profiles. However, improving and refactoring this to reflect operations in real time and to output more useful data will be part of the work of developing the new ingestion tools.

- **Data input formats and protocols:** With the goal of harmonizing metadata processing workflows between The European Library and the Europeana Foundation, the new system will firstly be developed to import XML data. The input schema for imported data in the system will initially consist of the most common library standards as well as ESE and EDM. As for the ingest protocols, FTP, HTTP and OAI-PMH will be supported and other protocols may be investigated. The Europeana Foundation will also need to make the upload of content and metadata in the Europeana Cloud as easy as possible.
- **Main metadata model used within the system:** In order to guarantee better interoperability of data and tools, The European Library will adopt EDM following the recommendations defined in the EDM profile for libraries⁵⁰ developed during the Europeana Libraries project. These recommendations will be reviewed in consultation with library data experts and may require the definition of new extensions of EDM. The Europeana Foundation will keep EDM as its main standard.
- **Data operations:** as described above, both The European Library and the Europeana Foundation perform harvesting, mapping and various enrichments on the data they aggregate. Building suitable processing tools for both is core to the development of the new system. The use cases and needs for mapping and enrichment are described below.

Mapping of aggregated data, from input data to EDM Internal data

Both The European Library and the Europeana Foundation data officers create and maintain specific mappings for each dataset. These are stored in their systems and applied with or without changes every time the dataset is re-ingested. Both also work with XSLT transformations, although in a different manner (see below). Retaining and being able to re-use and update the mappings is essential. As for the future implementation of a mapping and transformation service, the fact that The European Library and the Europeana Foundation share a common approach for that matter and the fact that the data officers who will be users of the new system have a similar technical background will inform the decisions made.

- While library domain expertise is highly needed for the mapping of library data to EDM, the fact that the number of input schema will be limited to the most common ones will help automate the generation of the transformed data.

⁴⁷ <http://pro.europeana.eu/share-your-data/data-guidelines/edm-profiles#Dataset>

⁴⁸ <http://labs.europeana.eu/api/provider>

⁴⁹ <http://www.europeana-inside.eu/home/index.html>

⁵⁰ <http://pro.europeana.eu/share-your-data/data-guidelines/edm-case-studies/europeana-libraries-edm>

- On the Europeana Foundation's side, automating fully the transformation of data to EDM Internal is desired.
- In the meantime, possibilities for analysing the data and updating the mapping files will be needed in both cases, in line with the type of customization performed until now, in order to improve the quality of the transformations.
- Editing data mapping files as described above will be done using a user interface consisting of an XSLT editor. The overall effort for the design of this user interface will focus on facilitating the analysis of the data to be mapped and previewing intermediary and final results. Some knowledge of XSLT may be required from the data officers.

An improved enrichment service

The Evaluation and Enrichments Task Force, conducted in 2015, highlighted crucial recommendations for the way data enrichments can be set up and processed.⁵¹ Both The European Library and the Europeana Foundation have already built a semantic enrichment service suited to their needs. Creating a shared customizable enrichment service that will enable enrichments, with appropriate targets and rules depending on the type of source data, will be a priority,⁵² as will working towards better quality of data and data enrichments. For instance, generating enrichments using both specific targets suitable for library data such as subject headings (e.g. Multilingual Access to Subjects (MACS))⁵³ as well as more generic targets suitable for the full Europeana dataset will be a requirement.

An important part of the work around enrichment will be to support the implementation of the Europeana Entities API and Europeana Collections creator pages and concept pages based on the API.⁵⁴

5.2.2 Development of Metis

Building a shared new set of tools dedicated to the aggregation of library data and publication of all Europeana EDM data, with the vision of a shared collaborative infrastructure in mind, is an opportunity to review and solve the drawbacks of our current systems. As identified while consulting with the Europeana DSI aggregating partners, there is a major need for a smoother aggregation process, in which data officers can focus on data quality more efficiently, and in which the outcomes of their work (validation feedback, data statistics, data enrichments) can be looped back to the institutions who contributed the data.

⁵¹ <http://pro.europeana.eu/taskforce/evaluation-and-enrichments>

⁵² See the recommendations for selecting target datasets for semantic enrichment:
http://pro.europeana.eu/files/Europeana_Professional/EuropeanaTech/EuropeanaTech_taskforces/Enrichment_Evaluation//EvaluationEnrichment_SelectingDatasets_102015.pdf

⁵³ http://pro.europeana.eu/files/Europeana_Professional/EuropeanaTech/EuropeanaTech_taskforces/Enrichment_Evaluation//AppendixA_InventoryEnrichmentTools_102015.pdf

⁵⁴ See Europeana DSI D6.2: Requirements for Europeana.eu,
http://pro.europeana.eu/files/Europeana_Professional/Projects/Project_list/Europeana_DSI/Deliverables/europeana-dsi-d6.2-requirementsforeuropeana.pdf

The system to be built, dubbed Metis,⁵⁵ can be defined as follows:

- Metis consists of several back-end services and a client (that together make up the product).
- The user interface of Metis offers services for harvesting, analysing, transforming, enriching and publishing data on the Europeana platforms.
- The first iteration of Metis targets the Europeana Foundation and The European Library data officers as its users. Some back-end services developed in the frame of Metis may be used by others; at a later stage Metis can be developed further and offered as a package of tools to expert hubs.
- Most of the services used by Metis are developed using cloud technologies; data processed in Metis is stored in the Europeana Cloud's storage layer. The Metis client itself is not necessarily hosted in the Europeana Cloud - it is one of its clients.

The back-end services developed as part of Metis are data processing services. Any data operation activated in the Metis user interface will be executed by one of the implemented data processing services, such as an EDM Validation service, a Data Statistics service, a Transformation service, an Enrichment service, a Dereferencing service,⁵⁶ and so on.

The services, initially designed to be used within Metis, will be developed in a modular manner with public APIs, enabling them to be used by other clients. As described, Metis for The European Library and the Europeana Foundation will use Europeana Cloud for storage, and all metadata and content will gradually be migrated there. (See the following sub-chapter [5.3](#)).

Experimenting with how data officers from an expert hub can coordinate their operations with the Europeana Foundation's data officers within a common system will enable better performance and reduce redundancy. Emphasis will be put on splitting aggregation tasks in an efficient way. For instance, actions such as mapping of data, instead of being performed twice - once by The European Library data officers and once by the Europeana Foundation data officers - adding overlap and unnecessary time for processing, will be triggered only by The European Library data officers. This effort, together with other improvements in the mechanisms for publishing to the Europeana platform and making the data officers' operations smoother will help solve the concerns highlighted by the Europeana Foundation's partners.

The development of Metis for the Europeana Foundation and The European Library will inform the possibility of extending the tool to serve an even wider range of domains and use cases. Such an extension will be part of Europeana DSI from 2017 onwards.

5.2.3 Technical design principles of Metis

With the long-term investment in the Europeana Cloud as a major aspect of the shared technical infrastructure in mind, a close collaboration between the team focused on building the Europeana

⁵⁵ After the Greek Titaness of Wisdom

⁵⁶ See <http://pro.europeana.eu/share-your-data/data-guidelines/europeana-semantic-enrichment#URI>

D1.1: WORK AND IMPLEMENTATION PLAN TO INNOVATE THE AGGREGATION INFRASTRUCTURE

Cloud storage service and the team developing Metis and its data services will be put in place. In December 2015, a workshop was organized in The Hague. The technical team from Work Package 2 of the Europeana Cloud project met with The European Library and technical members of the Europeana Foundation's Data Partner Services team to discuss recommendations and requirements that will enable and facilitate the aspects of Metis relying on Europeana Cloud for storage and parts of data processing. This collaboration will be pursued in 2016 in order to achieve the best performance level possible with the implementation of Metis.

The following principles will drive the development of Metis throughout the project:

- Distinct modular services will be designed to support data processing workflows. Where applicable, these services will rely on the Data Processing Service of Europeana Cloud. The resulting combination of services will be applied to datasets in predefined automated sequences.
- Some services and workflows will be common to The European Library and the Europeana Foundation; some will be shared with different customizations; some will remain specific to one group of users.
- All history of tasks performed on a given dataset will be stored locally in Metis and made accessible via relevant REST APIs.
- An effort will be made to keep, in addition to the tasks history, provenance information indicating which specific operations generated a new version of the processed data.
- Validation, preview and access to statistics on the data will be available at all times.
- Since the Metis orchestrating interface will initially be developed for two groups of users, special attention will be paid to an appropriate user (data officer) collaboration.
- Once Metis is judged capable of supporting the first two groups of users (The Europeana Foundation and The European Library), another round of user research and requirements analyses will be initialized, thus beginning the process of extending use to more user groups/aggregator partners.

The development of Metis as an aggregation client for The European Library and the Europeana Foundation will start in 2016 and will be pursued in 2017 up to the point where the previous version of the UIM can and will be deprecated. It is expected that in 2016 some of the planned Metis data processing services will also be made available to and re-usable within other Europeana DSI aggregators' systems (see following sub-chapter [5.3](#)).

The focus on improving the transparency of the aggregation processes within the Europeana Foundations' workflow as well as the effort put into making the processes smoother will benefit the Europeana Foundation's partners.

User research and requirements analyses for further aggregators to become Metis users will be initiated in the beginning of 2017 with earliest development provisionally planned for the second half of 2017.

5.3 Streamlining data quality related processes between the Europeana Foundation and Europeana DSI aggregators

As part of the services designed for Metis, technical services for validation of EDM data, quality statistics on a group of items described using EDM, and preview of EDM data in the Europeana Collections environment will be implemented and made available to partners in the form of REST APIs. Sharing these services from their prototyping phase will provide an opportunity to design them in collaboration with the technical DSI partners. Their implementation where possible in other existing systems should contribute to data quality improvements by standardizing the way EDM data is analysed. A first step towards this implementation will be taken as part of Europeana DSI-2 by Knowledge Integration, who will prototype the Europeana Inside Connection Kit (ECK) to use the planned EDM Validation service, the EDM Data Statistics service and the Europeana Collections Preview service.

The EDM Validation service and Data Statistics service will be designed to support the implementation of the Europeana Publishing Framework and gradually take into account, after setting the priorities, the recommendations of the Data Quality Committee initiated in January 2016. Both services will be implemented in parallel since it is believed that formal validation of data against a data model (including the provision of mandatory elements) is not sufficient information to evaluate the quality of data for a given scenario of use. As discussed in the EDM workshop in Amsterdam in November 2015:⁵⁷

“softer approaches to data checking could be designed. These would not necessarily focus on binary valid/invalid judgments, but rather on warning or outputting measures that show when the tested data fails to meet good standards for re-usability by a given user community, or only partially meets the requirements of a specific application.”⁵⁸

In August 2017, while Metis as a full suite with a user interface will only be in use for The European Library and the Europeana Foundation, a set of services, including the three mentioned above will have been documented and promoted so that other partners can investigate the possibility of using them independently to support better normalization of our practices.

⁵⁷ <http://pro.europeana.eu/event/edm-workshop-edm-turns-five-so-now-what>

⁵⁸ Valentine Charles, Antoine Isaac in <http://pro.europeana.eu/page/edm-turns-five-so-now-what-workshop>

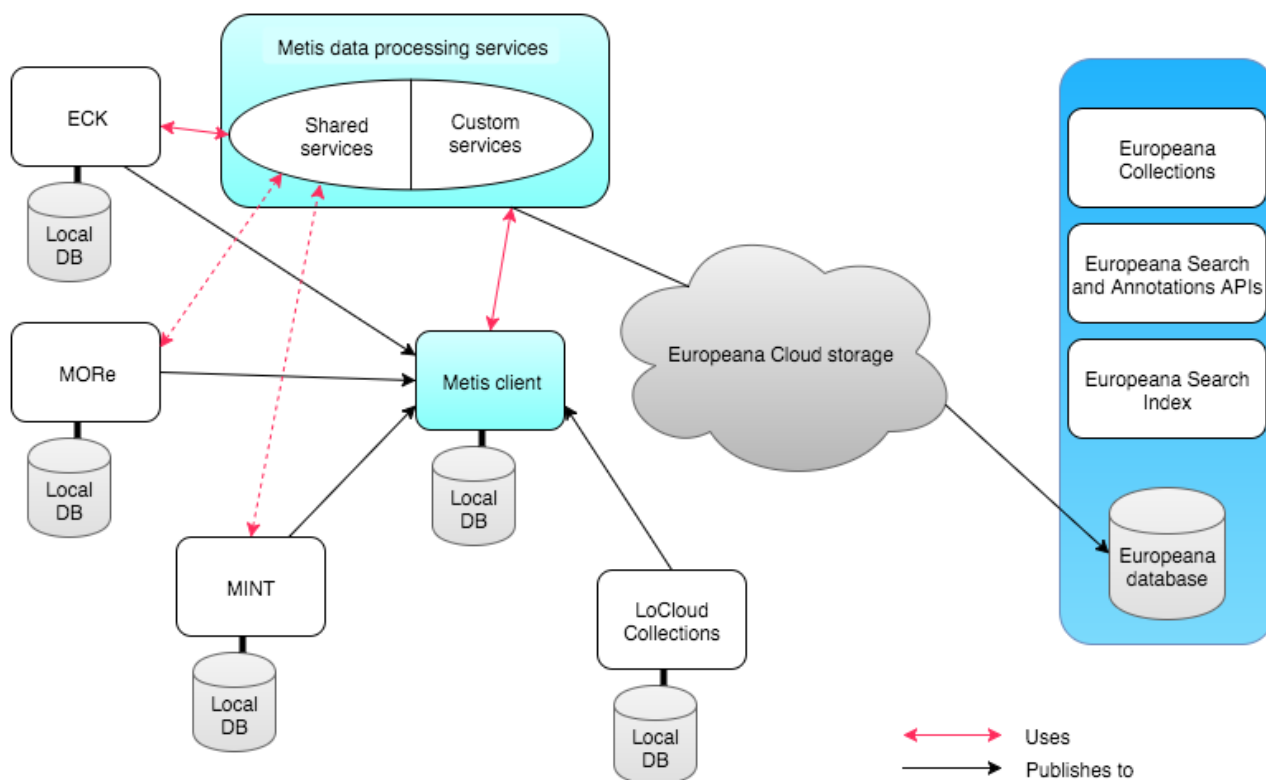


Figure 4. Schematic diagram representing the goal of Metis development by September 2017.

5.4 The future of Metis

5.4.1 Metis, a set of data ingestion tools to be used by all expert hubs

Once Metis is close to a first release for use by The European Library and the Europeana Foundation, we will begin to investigate the possibility of offering instances of Metis to support other expert hubs' operations. As already identified in the process of planning the development of Metis for The European Library and the Europeana Foundation, Metis as a set of data ingestion tools will focus exclusively on fully common or customizable workflows, and will have data officers for aggregation as its primary group of users. Provided that data officers from a given expert hub have the same user profile and similar use cases as the data officers Metis was initially developed for, using Metis as a tool shared by the Europeana Foundation and expert hubs is an option. This must be evaluated case by case.

Designing and developing Metis to cover the needs, use cases and users of all expert hubs will significantly increase the scope and complexity of the product. Extending the functionality of Metis to cover the needs of all groups of users described for the distributed approach is not recommended as the increase in functional scope would risk being overshadowed by architectural complexity and radically reduced maintainability, resulting in significant increases in development and operations.

An alternative strategy to designing and developing one tool to fit all use cases would be to adapt existing tools that are already offering very specific services so that they become interoperable with Metis. In this sense, future plans should focus on identifying what Metis as a set of tools can incorporate and what should remain external to Metis. For instance, the data mapping XSLT editor as envisioned for the first iteration of Metis will meet the needs of data officers who have basic knowledge of XSLT, but it will not be usable by an extended group of users with widely varying data expertise and backgrounds. On the other hand, the MINT platform, which has been used over time by many of the Europeana DSI aggregators⁵⁹ and their data providers to map and perform transformations, already offers a high level of user-friendly functionality as well as possibilities to map data from a wider range of formats.

Building an automated connection mechanism from such systems (MORe is another example) to Metis may be of greater value than working towards full consolidation of Metis and may better support DSI partners' specific areas of expertise.

5.4.2 Metis, a set of data services connected to external systems

Considering a shared infrastructure for cultural heritage data aggregation in Europe that would consist of a distributed infrastructure rather than a centralized set of tools is more in line with the network approach that the Europeana Foundation and its partners share. While a challenging amount of work on standardization will need to be pushed forward to support this approach, it is probable that this is a suitable way to benefit from the richness and diversity of our practices.

For this shared infrastructure to be in place, the two following directions will be investigated further:

- Shared services: While tools themselves are not necessarily common to all, some data services should be. It is recommended that the type of prototyping described in the previous sub-chapter [5.3](#) is extended to encompass additional partners and systems at a later stage.
- Connected operations between platforms: Metis should support data exchange and delivery with existing aggregation platforms and data management applications.

5.5 Beyond Metis

As outlined in 3.4 above, research and development will be initiated in the second half of 2016 with the aim of identifying, testing and prototyping successor solutions to Metis. The proposed solutions should support radically different ways, both organizational and technological, of fulfilling the Europeana mission.

⁵⁹ Europeana Fashion, CARARE, Photoconsortium, MICHAEL, EUscreen.

6 A content strategy for the Europeana DSI

6.1 Development of the Europeana database over time

Europeana started in 2008 with the “aim of sharing Europe’s culture and heritage in the online world”.⁶⁰ A critical mass of content was built between 2008 and 2012, based on a simple data model, the Europeana Semantic Elements (ESE), and low minimum data quality standards. This strategy allowed the Europeana database to grow fast, but has left a legacy of very heterogenous metadata records in the Europeana database.

In 2012, Europeana released all metadata into the public domain by using the CC0 Public Domain Dedication and has since then published metadata as CC0 under the Europeana Data Exchange Agreement.⁶¹

In 2013, Europeana migrated the data to the Europeana Data Model (EDM), a model that allows a richer and more nuanced representation of the data.

In November 2015, about 50% of the records in the Europeana database were provided in native EDM, which still leaves a legacy of about 20 million records in EDM automatically converted from ESE. With the shift to EDM, rights information became mandatory and since then has been uniformly applied across the Europeana database.

Beyond the above changes, the minimum quality criteria (e.g. mandatory fields or normalization of field values) have not been increased. Despite data quality being a stated goal for the last few years, data quality overall is still too low and the corpus as a whole is highly heterogeneous. Quantitative growth has continued to be significant and even increased in pace. The average monthly record increase rates for 2012, 2013 and 2014 were 144,947, 581,406 and 725,216 respectively. The average increase for 2015 is 756,143 records a month. Were we to grow at the same pace in the coming years there would be more than 100 million records in the Europeana database by 2020.⁶² Without significant improvements to quality, the visitor experience of Europeana Collections will deteriorate as discovery of quality content will become increasingly difficult for all types of users.

Since its inception, Europeana platform has been supply driven with data being provided via a network of aggregators, and the Europeana Foundation ingesting and publishing what has been offered. Although a number of EU-funded projects have been built around assumed existing content gaps to strengthen certain domains or themes in the Europeana database, there was no systematic feedback loop established connecting user and market demand to supply.

⁶⁰ WP1 EuropeanaNet, Jill Cousins, Harry Verwayen, Mel Collier: Outline Business Plan for Europeana as a service of the EDL Foundation, Nov 2008, p9

⁶¹ <http://pro.europeana.eu/page/the-data-exchange-agreement>

⁶² see Europeana Data Ingestion Plan:

http://pro.europeana.eu/files/Europeana_Professional/Projects/Project_list/Europeana_DSI/Milestones/europeana-dsi-ms6-data-ingestion-plan.pdf

For the reasons listed above, Europeana Foundation aggregators and other members of the Europeana Network Association have been re-thinking the approach to collecting data. This resulted in the *Europeana Publishing Framework* which was released by the end of 2015. The framework clearly defines what the Europeana Foundation and partners can do with cultural heritage data depending on the technical qualities and copyright status of its linked media. Every record in the Europeana database can now be matched to one of four tiers, with each consecutive tier having increasingly higher requirements. The higher the tier, the more can be done with the material such as feature it in one of the Europeana Thematic Collections or on Europeana Labs.

The framework also gives guidance on how to identify new data and improve existing data by actively identifying gaps and opportunities and work with both existing and potential new partners to fill them. In Europeana's business plan for 2015, three themes were set: art history, music and fashion. We will continue to actively look for collections that fit these themes and put extra effort into making them available in the best way possible.

The consequence of this is that data that does not fit the themes set for that year, or that qualifies into a lower tier, will be lower priority than data that matches the themes and qualifies for higher tiers. All data that conforms to the mandatory metadata elements and the minimum content requirements will still be accepted but will be explicitly prioritized based on tier and theme rather than the current implicit first come, first served policy.

6.2 Towards a content strategy

Content strategy can be defined as "Getting the right content to the right user at the right time through strategic planning of content creation, delivery, and governance."⁶³ To improve the capabilities of the Europeana DSI to deliver on the *Europeana Strategy 2015-2020*⁶⁴ and on the annual business plans, a *Europeana Content Strategy* will be developed.

A fundamental principle of the content strategy will be that user demand will be the major factor in shaping content acquisition, content improvements and quality criteria ("right content to the right user"). A major aspect of the content strategy will thus be to outline the methodological approach the Europeana DSI will adopt in researching the demand that target user groups have on content.

The other aspects to a content strategy - content creation, content delivery and content governance - will also be covered.

Other more specific aspects likely to be covered in the content strategy are:

- Technical quality criteria for metadata and digital objects
- Automated data enrichments
- Multilingual discovery from a data perspective
- The role of user contributions

⁶³ <http://contentstrategyalliance.com/the-beginnings/csa-charter/>

⁶⁴ <http://strategy2020.europeana.eu/>

As the content strategy covers a long time period, and user demand is constantly changing and evolving, it will focus on processes and policies rather than on setting specific criteria or rules.

The content strategy will be drafted in 2016 by a cross-departmental team of the Europeana Foundation. The cross-departmental team will consult with the newly established Data Quality Committee in regards to defining a metadata quality framework for Europeana to complement the rights and media-centric *Europeana Publishing Framework*. The Data Quality Committee is a standing working group that reviews the Europeana DSI data quality standards and develops them over time.

It is foreseen that the content strategy will be ratified by the Aggregators' Forum and the Europeana Collections and Europeana Labs editorial teams later in 2016. The content strategy will be reviewed annually to discuss necessary updates. These reviews will be based on and closely follow the annual business plans of the Europeana Foundation.

7 Roadmap and conclusion

This roadmap will mainly look at what is planned to happen between January 2016 and mid-2017. Further refinements are planned for the time when the *Europeana Business Plan 2017* is being written. A visual timeline is provided in Figure 5, further explanations are added below.

With regards to the shared infrastructure development, three areas of work will be addressed in the first half of 2016.

1. The EDM Validation service will be prototyped as a standalone service to make sure any records submitted to the service can be validated against the EDM schema and the mandatory EDM elements. Data partners will then have the chance to validate their records before submitting them to the Europeana DSI.
2. The functional requirements from The European Library and the Europeana Foundation for Metis will be finalized in order to prepare for the actual development of the toolset.
3. The UX design for the Metis interface is also scheduled for the first half of 2016.

In late spring 2016, we will start to build on the outcomes of the Europeana Cloud project and adapt the Europeana Cloud infrastructure to support the development of the back-end services. In parallel, the Data Statistics and Collections Preview services will be developed. We expect these two services to be available for other aggregators to use and build upon by the end of 2016. The Metis interface development will also start in mid-2016 and the first back-end services will be connected to it to see how these services will work together.

Before each of the three mentioned back-end services become publicly available, technical partners can start prototyping for their tools to make use of these services. An example of these tools is the Europeana Connection Kit (ECK): we will consult with the technical partner during the development of the services, evaluate the possibilities for building on previous experiences and eventually work on prototyping the ECK to use the EDM Validation service, EDM Data Statistics service and Europeana Collections preview service. We will consult with more technical partners of course and hope for more prototypes to be available.

D1.1: WORK AND IMPLEMENTATION PLAN TO INNOVATE THE AGGREGATION INFRASTRUCTURE

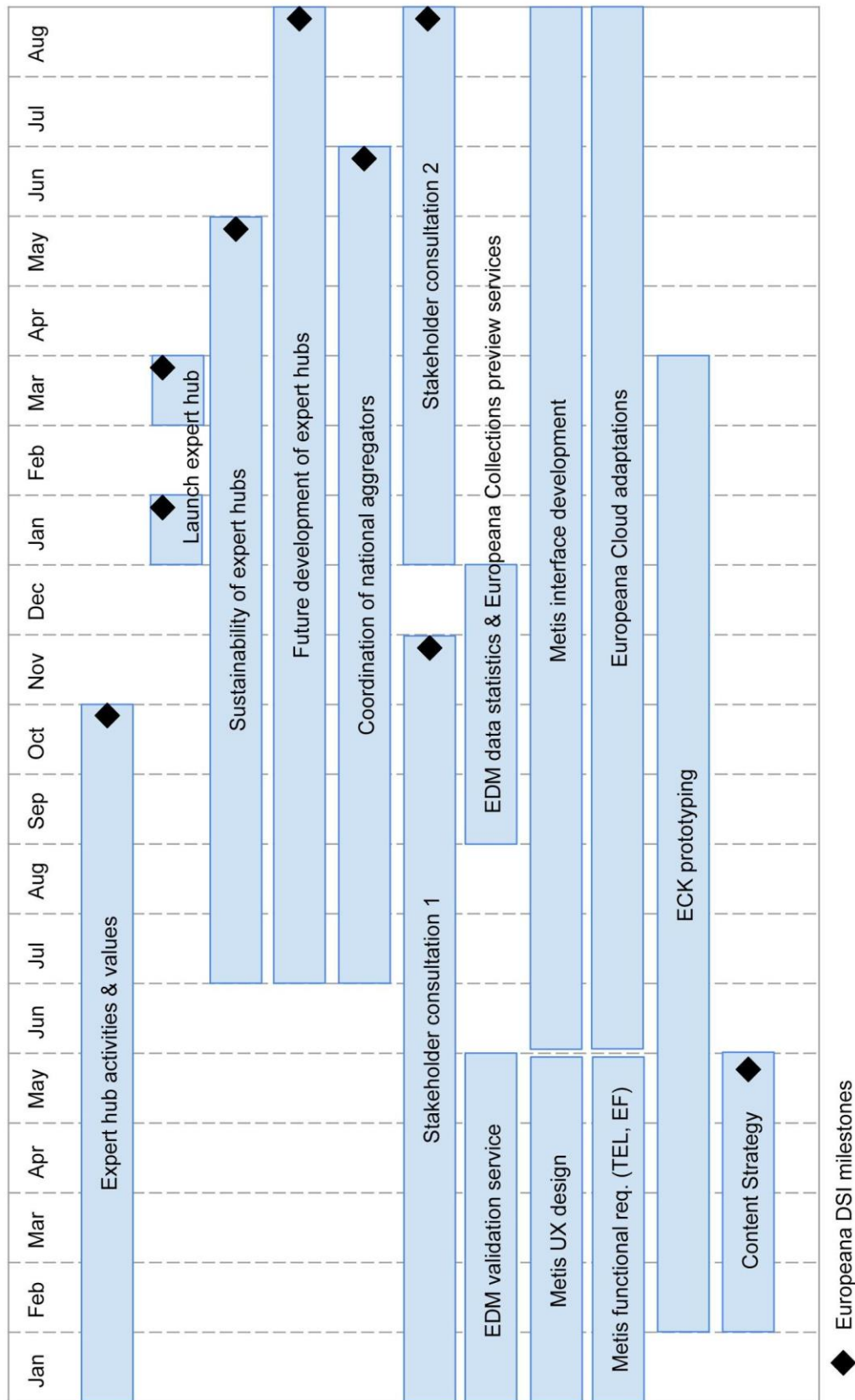


Figure 5. Aggregation infrastructure development roadmap, covering the period from Jan 2016 to Aug 2017. Note that it excludes the long-term research and development activities mentioned in chapter 3.4.

D1.1: WORK AND IMPLEMENTATION PLAN TO INNOVATE THE AGGREGATION INFRASTRUCTURE

In parallel to the work described above, we will work with aggregators on their requirements for a shared aggregation infrastructure. This work will be ongoing until August 2017. While we have done this recently in workshops with a group of aggregators, we propose to continue in one-on-one settings: working with one aggregator at a time to understand the specific needs of each aggregator and respond to them appropriately. We will also use the bi-annual Aggregators' Forum meetings to follow up with the group of aggregators and provide an update on the progress and next steps.

In parallel, we are going to further develop the expert hub concept and launch the first expert hubs in 2017. We will also continue the dialogue of existing and future expert hubs and look into the first lessons learned of the expert hubs being launched. In 2017, we will also create a sustainability plan with domain aggregators who are becoming expert hubs. Using the work of the Europeana Network Association Task Force led by Europeana Fashion, a blueprint for the sustainability of the expert hubs will be drawn up by the expert hub candidates in consultation with the Europeana Foundation.

We will also continue to look beyond domain aggregators and work with national aggregators on their role in the aggregation landscape. We will also develop a joint plan on how best to manage the exchange of data between national aggregators, expert hubs and the Europeana Foundation.

From the end of 2017 onwards, we expect to be ready to launch more expert hubs and extend Metis to meet the requirements of these expert hubs. At this stage of the development process, it is quite unlikely that Metis will fully support a distributed workflow in the same sense as MINT does now for data partners of aggregators. However, Metis will support a full ingestion and publication pipeline that we can offer to aggregators to process their data efficiently from source to Europeana Collections. It will also collaborate with existing tools, so aggregators can continue using their own tools while reaping the benefits of Metis. In this way, these external tools or sets of tools become part of the shared aggregation infrastructure that is available to expert hubs to make data ingestion and publication as smooth as possible.

Data officers working for the expert hubs will be able to publish directly to Europeana Collections without an extra layer of Europeana Foundation data officers doing this for them. This alone will improve efficiency and allow data to be published, updated or removed more quickly. It also improves efficiency for the Europeana Foundation data officers as they will have the time to look more closely into data partners not using Metis or data partners submitting data that we requested to be provided to support a thematic collection, for example.

Even this is not the end of the story. In parallel to all the work listed above, a group of experts will explore completely different data acquisition mechanisms. They may come up with even more innovative or revolutionary scenarios or technologies we can use to make unlocking Europe's digital cultural heritage even easier and to surface the highest possible quality of digital cultural heritage.

Appendix

Definitions

- **Aggregator:** An organization that collects, formats and manages metadata from multiple Data Providers, providing services such as offering their own portal and acting as Data Provider to Europeana.⁶⁵
- **Data officer:** A person who in a professional capacity harvests, maps, enriches and transforms data on behalf of data partners for the ultimate purpose of making the data available on the Europeana platform. Note that this is a role, not a job title.
- **Europeana Collections:** The main website to discover the vast and amazing digitized collections of Europe's cultural heritage institutions.⁶⁶
- **Europeana Digital Service Infrastructure (Europeana DSI):** Funded by the Connecting Europe Facility (CEF), the Europeana Foundation and all aggregators are building the 'Europeana Platform', which is the DSI. The core objectives of the Europeana DSI are to innovate the aggregation infrastructure, boost the distribution infrastructure and work towards long-term financial stability through business model innovation.
- **Europeana Foundation:** The Europeana Foundation is the governing body of the Europeana platform. It has about 60 staff, based mainly in The Hague. It reports to a 17-strong Governing Board of representatives from professional associations of cultural and scientific heritage organizations, who advise on policy and strategy.
- **Europeana Publishing Framework:** It clarifies the relationship between the Europeana DSI and data partners. It clearly outlines what the Europeana DSI will do for data partners depending on the content and metadata data partners decide to make available.⁶⁷
- **Europeana Thematic Collections:** Curated collections on popular interest topics; Europeana Music Collections and Europeana Art History Collections are the first two collections that will launch officially in early 2016.

⁶⁵ Taken from the Europeana Glossary of Terms: <http://pro.europeana.eu/page/glossary>.

⁶⁶ <http://www.europeana.eu/>

⁶⁷ <http://pro.europeana.eu/publication/publishing-framework>