Europeana – Core Service Platform

MILESTONE

MS6: Data ingestion plan

Revision	1.0
Date of submission	30 October 2015
Author(s)	Jeroen Cichy (EF), Henning Scholz (EF), Kerstin Herlt (ACE), Julia Welter (DIF), Teresa Natale (MICHAEL), Marzia Piccininno (MICHAEL), Gisela Baumann (FUB), Kate Fernie (2Culture), Marco Rendina (eFashion), Johan Oomen (NISV), Andrew Ormsby (BUFVC), Afelonne Doek (IALHI), Gariella Ivacs (IALHI), Adina Ciocoiu (TEL), Anastasia Gasia (TEL)
Dissemination Level	Public

Co-financed by the European Union
Connecting Europe Facility

REVISION HISTORY AND STATEMENT OF ORIGINALITY

Revision History

Revision No.	Date	Author	Description
0.0	30/09/2015	Jeroen Cichy (EF)	Initial draft
0.1	09/10/2015	Jeroen Cichy (EF), Henning Scholz (EF)	Second draft, comments integrated
0.2	20/10/2015	Jeroen Cichy (EF), Henning Scholz (EF), Kerstin Herlt (ACE), Julia Welter (DIF), Teresa Natale (MICHAEL), Marzia Piccininno (MICHAEL), Gisela Baumann (FUB), Kate Fernie (2Culture), Marco Rendina (eFashion), Johan Oomen (NISV), Andrew Ormsby (BUFVC), Afelonne Doek (IALHI), Gariella Ivacs (IALHI), Adina Ciocoiu (TEL), Anastasia Gasia (TEL)	Inclusion of all ingestion plans for the DSI partners, comments integrated
1.0	30/10/2015	Jeroen Cichy (EF), Henning Scholz (EF)	Final comments integrated, layout edits, minor changes

Statement of originality:

This milestone contains original unpublished work except where clearly indicated otherwise.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. Introduction

1.1 What is ingestion?

According to the definition from the Europeana Glossary of Terms[1] ingestion is: "The process of collecting, mapping and publishing the data from the Data Provider to Europeana." On a more basic level this means updating and inserting metadata records into the Europeana database and making those records available in the europeana.eu. This document will not go into the details of the entire ingestion workflow, but will focus more on the data to be ingested. In short, the ingestion workflow for Europeana consists of harvesting metadata from an aggregator, validating the metadata and mapping it to EDM, enriching the metadata and publishing it. For more information on the ingestion workflow see the Europeana Publishing Guide[2] (formerly known as the Europeana Publication Policy).

1.2 Purpose of the Ingestion Plan

The ingestion plan will make a planning for metadata to be ingested by Europeana. It is to some extent complementary to the DSI deliverable D1.3, Description of planned work for aggregators on improving data. D1.3 is not published yet, but will have a plan how each DSI partner is improving data quality, which will result in updates of existing datasets. The ingestion plan focuses on new submissions of data, unless stated otherwise.

A division has been made between the different sources of the metadata to be ingested. The first chapter will deal with the ingestion for the DSI partners for the first year of funding of the Digital Service Infrastructure (DSI). The second chapter will look at the ingestion for the EU funded projects that are still ongoing. A planning will be made for the duration of the project. The third chapter will try to make an estimation on the expected amount of metadata to be ingested for national aggregators and other data partners.

1.3 Data partners

The partners delivering metadata to Europeana will for the purpose of this document be divided into three main categories:

1. DSI partners (domain and thematic aggregators);
2. EU funded projects;
3. Other partners (national aggregators and direct providers).

The first type concerns the full partners of the Europeana Digital Service Infrastructure (DSI) and are funded by the European Commission for the duration of the DSI. For the DSI partners the scope of the ingestion plan will be the first year of the DSI, starting on 1 April 2015 until 31 May 2016.

EU funded projects receive funds for the duration of a project. For them the scope will be 1 October 2015 until the end of the project, which will vary for the different projects.

---

[1] Glossary of Terms: http://pro.europeana.eu/page/glossary

[2] http://pro.europeana.eu/publication/publication-policy

The remaining partners consist of national aggregators and some direct providers, i.e. institutions that do not act as an aggregator, but provide metadata to Europeana directly. For them the scope of the ingestion plan will be 1 October 2015 until 31 December 2015.

A chapter will be dedicated to each of these partner categories.

1.4 Key performance indicators 2015

For Europeana some key performance indicators (KPI's) related to the ingestion of metadata for the end of 2015 have been drawn up, as published in the Business Plan 2015[3]. They are as follows:

- 45 million metadata records accessible in Europeana;
- 25 million metadata records provided in native EDM;
- 18 million digital objects labelled under rights statements that allow re-use (PDM, Creative Commons, OOC-NC);
- 13 million downloadable digital objects with direct links;
- 18 million digital objects include geo-location data.

The progress of these KPI's is monitored in the form of a monthly content report, which is created by one of the members of the Data Partner Services Team. The table below shows the current status as of 13-7-2015[4] compared to the KPI's.

Table 1: KPI's compared to actual metadata records

	KPI end 2015	July 2015	On track
Metadata records in Europeana	45,000,000	44,725,949	🟢
Downloadable digital objects with direct links	13,000,000	Between **12** and **25 million** depending on method of calculation	🟢
Number of metadata records provided in EDM	25,000,000	26,601,956	🟢
Increase in number of digital objects with geo-location data	18,000,000	12,979,216	🟡
Number of digital objects labelled under rights statements that allow re-use (PDM, CC0, Creative Commons, OOC-NC)	18,000,000	22,163,992	🟢

---

[3] See Annex 2 of the Business Plan 2015:
http://pro.europeana.eu/files/Europeana_Professional/Publications/Make%20the%20Beautiful%20Thing%20Business%20Plan%202015.pdf
[4] This is the most recent publication to the Europeana portal.

Figure 1 is an updated version of the graph published in the Partner and Data Development Plan[5] and shows the growth of the amount of records in the Europeana portal over the last five years. As can be seen the average rate of increase has increased over time. For comparison, the average monthly increase rates for 2012, 2013 and 2014 were 144,947, 581,406 and 725,216 records, respectively. The average increase for 2015 (1 Januari - 13 July) currently is 908,178 records a month. As can be shown, the objective of 45 million metadata records has almost been achieved and is expected to be greatly surpassed by the end of the year. If we are assuming a similar increase in records over the next years, we will have about 55 million records by the end of 2016 and about 65 million records by the end of 2017. We should be able to reach 100 million records by 2020, if we continue to be inclusive to all digital cultural heritage. A more quality based approach could slow down the growth rates in the next years. It is therefore extremely important that we keep on monitoring the progress on the (speed of) growth of the Europeana dataset. Related to this Europeana is currently working on a content strategy to manage this increase and move to an improvement of quality rather than quantity.
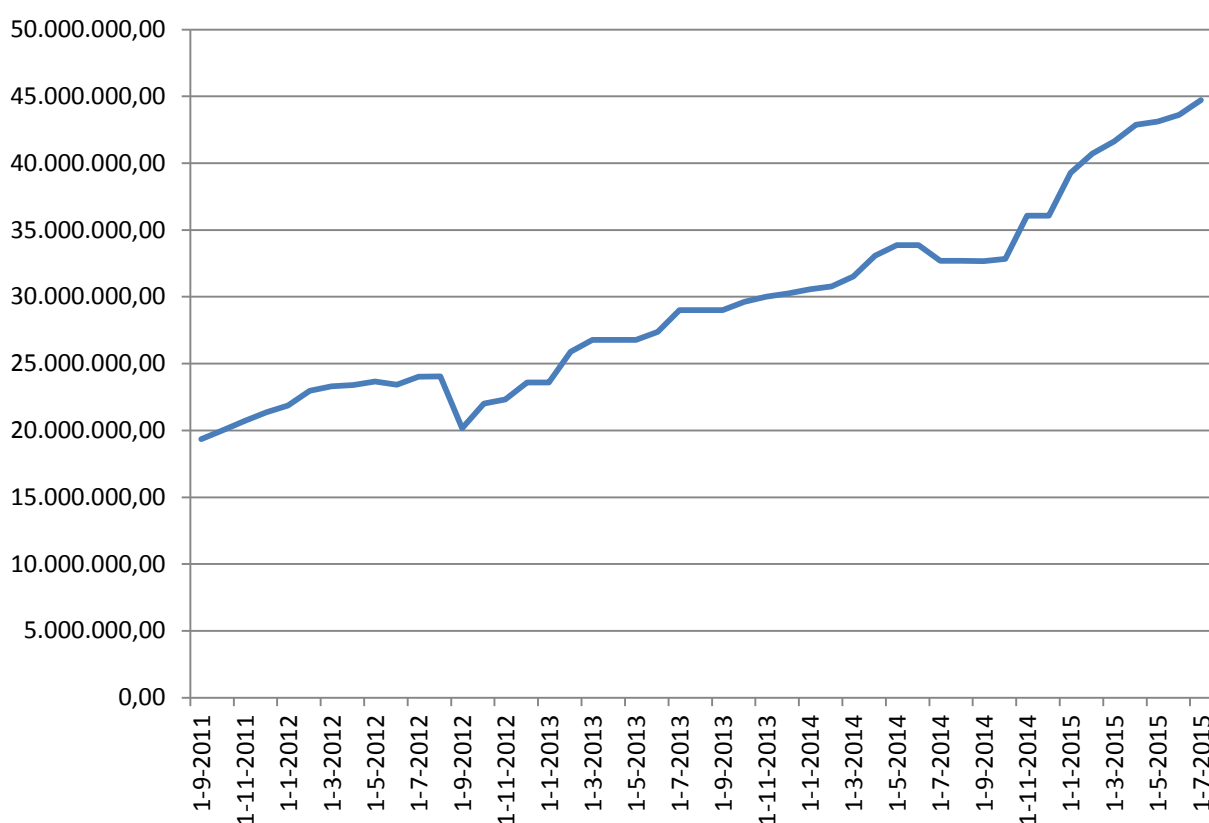
Figure 1: Number of records in Europeana from 9-2011 to 7-2015

1.5 Ingestion objectives

The objectives for the ingestion of new metadata have been described in the Partner and Data Development Plan (MS2, MS4 and MS5)[6]. As Europeana is becoming a platform, where the re-

---

[5] Partner and Data Development Plan, figure 1:
http://pro.europeana.eu/files/Europeana_Professional/Projects/Project_list/Europeana_Version3/Milestones/Ev3%20MS2%20MS4%20MS5%20Partner%20and%20Data%20Development.pdf
[6]

http://pro.europeana.eu/files/Europeana_Professional/Projects/Project_list/Europeana_Version3/Milestones/Ev3%20MS2%20MS4%20MS5%20Partner%20and%20Data%20Development.pdf

use of high quality material is desired, the overall goal is to improve the data quality and focus less on the quantitative growth of the Europeana database. To achieve this, there will be a focus on the accessibility, accuracy, consistency and re-usability of the data. Therefore, the desired metadata to be ingested will preferably:

- have previews of the cultural heritage objects;
- have direct links to the cultural heritage objects;
- be delivered in native EDM;
- have rights labels that allow re-use (Public Domain Mark, Creative Commons or Out of Copyright – Non-commercial re-use).

These objectives are also reflected in the KPI's (see paragraph 1.5).

2. Ingestion for DSI partners

This chapter will focus on the planned ingestion for the first year of the DSI as provided by each of the DSI aggregating partners. Table 2 shows an estimation of the total amount of new metadata records to be ingested by Europeana per DSI partner. The actual number of records might change as the numbers for some sets still need to be confirmed and for some sets the numbers are rough estimates. Where possible a distinction was made between new metadata from updated sets and metadata from new sets. Updates have been marked in light grey in the tables of the individual providers.

Table 2: Ingestion totals for each of the DSI partners

DSI partner	Metadata is part of updated set	Metadata is part of new set	Total
European Film Gateway	not specified	not specified	6,893
MUSEU	not specified	not specified	66,080
OpenUp!	5,600	347,000	352,600
CARARE	30,860	9,202	40,062
EuropeanaFashion	not specified	not specified	52,467
EUscreen	not specified	not specified	915,486
HOPE2.0 / IALHI	not specified	not specified	55,035
The European Library	154,119	397,312	551,431
Total			**2,040,054**

2.1 European Film Gateway

Table 3: Planned ingestion for EFG

Month/Quarter	Data provider	Estimated number of records	Type	Subject	IPR (content)
Oct - Nov 2015	Cineteca Nazionale	9	Video	Italian non-fiction films from WWI	Orphan works
Oct - Nov 2015	Svenska Filminstitutet	350	Text	Programmes of Swedish silent films	Rights Reserved - Free Access; partly PD
Dec 2015 – Feb 2016	Deutsches Filminstitut	250	Video	German films: Trailers, short films, excerpts	Rights Reserved - Free Access
Dec 2015 – Feb 2016	Deutsches Filminstitut	5,000	Image	Film stills from German films	Rights Reserved - Free Access
Dec 2015 – Feb 2016	Filmoteka Narodowa	400	Video	Pre-war newsreels of the state-owned Polish news agency Polska Agenja Telegraficzna (1924–1939)	Rights Reserved - Free Access
Mar – May 2016	Arhiva Naţională de Filme	6	Video	Early Romanian films (scientific, travelogue, reportage, feature)	Rights Reserved - Free Access
Mar – May 2016	Cinemateca Portuguesa	120	Video	Portuguese silent films	Partly Rights reserved, partly PD

Mar – May 2016	Cinematheque of Macedonia	16	Video	Films from Macedonian film pioneers Manaki Brothers + 8 Macedonian feature films from 1960-1990	Rights Reserved - Free Access
Mar – May 2016	Filmoteca de Catalunya	150	Video	Films on the Spanish Civil War produced by Laya Film, films by Segundo de Chomón, Spanish film pioneer (productions from 1903-1912)	Rights Reserved - Free Access, partly PD
Mar – May 2016	Fondazione Cineteca di Bologna	22	Video	Balkan War and WWI	Rights Reserved - Free Access, partly orphan
Mar – May 2016	Fondazione Cineteca di Bologna	170	Image	WWI, Bologna	Rights Reserved - Free Access, partly orphan
Mar – May 2016	Museo Nazionale del Cinema di Torino	150	Video	Rare Italian fiction and non-fiction films from the silent era from 1900-early 1920s + amateur films from 1920s to 1960s showing social and urban changes of the city of Torino	Rights Reserved - Free Access, PD and orphans
Mar – May 2016	Österreichisches Filmmuseum	250	Video	Austrofascist newsreel series "Österreich in Bild und Ton" from 1935-1937	Rights Reserved - Free Access
	Total	**6,893**			

2.2 MUSEU

Table 4: Planned ingestion for MUSEU

Month/Quarter	Data provider	Estimated number of records	Type	Subject	IPR (content)
Dec 2015	Royal Collection	150	Image	Art, decorative art, armour, music, and photographs	Rights Reserved - Free Access
Jan 2016	Telegraph Museum	100	Image	Collection related to telegraph technology.	Rights Reserved - Free Access
May 2016	Jersey Heritage	10,000	Image	Mainly coin hoards	Rights Reserved - Free Access
May 2016	Portable Antiquities Scheme	30,000	Image	Finds, of all dates, made by members of the public in the UK	Rights Reserved - Free Access
May 2016	Fitzwilliam Museum	15,000	Image	Wide range of subjects	Rights Reserved - Free Access
Q1 2016	Internet Culturale	10,000	Sound	Italian music collection	Rights Reserved - Free Access
Q1 2016	Regione Umbria	200	Image	Images of various museums of Regione Umbria	Rights Reserved - Free Access
Q1 2016	Internet Culturale	200	Text	New acquisitions from the Italian libraries	Rights Reserved - Free Access
Q1 2016	Biblioteca Nazionale Centrale di Roma	80	Image	Van Vittel's landscapes	Rights Reserved - Free Access
Q1 2016	GoTellGo	150	Image	Street art	Rights Reserved - Free Access
Q1 2016	Biblioteca Vaticana	100	Text	Manuscripts	Rights Reserved - Free Access
Q2 2016	Museo Italiano della Ghisa	100	Image	Pictures of XIX and XX centuries artwork made of cast iron	Rights Reserved - Free Access
	Total	**66,080**			

2.3 OpenUp!

Table 5: Planned ingestion for OpenUp!

Month/Quarter	Data provider	Estimated number of records	Type	Subject	IPR (content)
Q4 2015	Muséum national d'Histoire naturelle	ca. 200	Image	Preserved plants and animals	CC BY-NC-ND
Q4 2015	Royal Botanic Gardens, Kew	ca. 200	Image	Preserved plants and economic botany specimens	Rights Reserved - Free Access /CC-BY-NC-SA
Q4 2015	The Royal Botanic Garden Edinburgh	ca. 200	Image	Living plant images and Herbarium images	CC-BY-NC-SA
Q4 2015	Natural History Museum, London	ca. 200	Image	Preserved plant and animal specimens	CC-BY
Q4 2015	Botanic Garden and Botanical Museum Berlin-Dahlem	ca. 200	Image	Preserved plant specimens	CC-BY-SA
Q4 2015	Naturalis Biodiversity Center	?	Image/Sound	Butterfly images, frog sounds	Rights Reserved - Free Access
Q4 2015	ETI BioInformatics, Leiden	0	Image/Sound/Video		Rights Reserved - Paid Access
Q4 2015	Biologiezentrum der Oberoesterreichischen Landesmuseen	ca. 200	Image	Preserved mosses, plants & animals	CC-BY-SA
Q4 2015	Botanic Garden Meise, Belgium	ca. 200	Image	Preserved plants	CC BY-NC-ND
Q4 2015	University of Tartu, Natural History Museum	ca. 200	Image	Preserved plants, mosses and butterflies	CC-BY-SA
Q4 2015	Museum für Naturkunde Berlin	ca. 200	Image	Protozoa and algal drawings, insect trays, preserved insects	CC-0 (insect trays); CC-BY-SA (drawings, insects, minerals)
Q4 2015	Museum of Geology, University of Tartu	ca. 200	Image	Fossils	CC-BY-NC
Q4 2015	Institute of Geology at TUT	ca. 200	Image	Microfossils, fossils	CC-BY-NC
Q4 2015	Museum fuer Naturkunde Berlin, Tierstimmenarchiv	ca. 200	Sound	Bird and other animal sounds	CC-BY-NC-SA
Q4 2015	Institute of Botany of Slovak Academy of Sciences	ca. 200	Image	Preserved plant specimens	CC-0
Q4 2015	Royal Belgian Institute of Natural Sciences	ca. 200	Image	Preserved Coleoptera (bugs)	CC-BY-SA
Q1 2016	University Museum of Bergen	100,000	Image	Preserved plant specimens	CC-BY
Q1 2016	UiO - Natural History Museum Oslo	200,000	Image	Preserved plant specimens	CC-BY
Q1 2016	Herbarium Erlangense	2,000	Image	Preserved plant specimens	CC-BY-NC-SA
Q1 2016	University of Tübingen	7,000	Image	Preserved plant specimens	CC-BY-SA
Q1 2016	Herbarium Haussknecht	28,000	Image	Preserved plant specimens	Rights Reserved - Free Access
Q1 2016	Herbarium Institut für Geobotanik und	10,000	Image	Preserved plant specimens	Rights Reserved - Free Access

Month/Quarter	Data provider	Estimated number of records	Type	Subject	IPR (content)
	Botanischer Garten, Martin-Luther-Universität Halle				
Q1 2016	Muséum national d'Histoire naturelle	ca. 200	Image	Preserved plants and animals	CC BY-NC-ND
Q1 2016	Royal Botanic Gardens, Kew	ca. 200	Image	Preserved plants and economic botany specimens	Rights Reserved - Free Access/CC-BY-NC-SA
Q1 2016	The Royal Botanic Garden Edinburgh	ca. 200	Image	Living plant images and Herbarium images	CC-BY-NC-SA
Q1 2016	Natural History Museum, London	ca. 200	Image	Preserved plant and animal specimens	CC-BY
Q1 2016	Botanic Garden and Botanical Museum Berlin-Dahlem	ca. 200	Image	Preserved plant specimens	CC-BY-SA
Q1 2016	Naturalis Biodiversity Center	?	Image/ Sound	Butterfly images, frog sounds	Rights Reserved - Free Access
Q1 2016	Biologiezentrum der Oberoesterreichischen	ca. 200	Image	Preserved mosses, plants & animals	CC-BY-SA
Q1 2016	Botanic Garden Meise, Belgium	ca. 200	Image	Preserved plants	CC BY-NC-ND
Q1 2016	University of Tartu, Natural History Museum	ca. 200	Image	Preserved plants, mosses and butterflies	CC-BY-SA
Q1 2016	Museum für Naturkunde Berlin	ca. 200	Image	Protozoa and algal drawings, insect trays, preserved insects	CC-0 (insect trays); CC-BY-SA (drawings, insects, minerals)
Q1 2016	Museum of Geology, University of Tartu	ca. 200	Image	Fossils	CC-BY-NC
Q1 2016	Institute of Geology at TUT	ca. 200	Image	Microfossils, fossils	CC-BY-NC
Q1 2016	Museum fuer Naturkunde Berlin, Tierstimmenarchiv	ca. 200	Sound	Bird and other animal sounds	CC-BY-NC-SA
Q1 2016	Institute of Botany of Slovak Academy of Sciences	ca. 200	Image	Preserved plant specimens	CC-0
Q1 2016	Royal Belgian Institute of Natural Sciences	ca. 200	Image	Preserved Coleoptera (bugs)	CC-BY-SA
	Total	**352,600**			

2.4 CARARE

Table 6: Planned ingestion for CARARE

Month/Quarter	Data provider	Estimated number of records	Type	Subject	IPR (content)
April 2015	Narodowy Instytut Dziedzictwa	59	Images	Archaeological monuments – for Google Field Trip pilot	CC-BY-SA
April 2015	Národní památkový ústav	update 12,577 (10,964 new)	Images, text	Archaeological monuments	Rights Reserved - Free Access
Sept 2015	KNAW-DANS	update 9,352 (2,896 new)	Images, text	Archaeology	Rights Reserved - Free Access
Q4 2015	Archaeology data service	Update (number of new records tbc) 34,197 existing	Images, text, 3D	Archaeology	Rights Reserved - Free Access
Q4 2015	Vilnius University Faculty of	21,190 update and enrichment, no	Images	Archaeology	Rights Reserved - Free Access

	Communication	new items			
Q4 2015	University of Jaen, Institute for Iberian Archaeology	Update to de-duplicate (estimated 1,207)	Images, video, text, 3D	Archaeology	CC-BY-NC (4563) CC-BY (303)
Q4 2015	Direction des Monuments et des Sites de la Région de Bruxelles-Capitale	13,590 update to fix links and display of thumbnails. Update planned in Q1 2016.	Images, text	Archaeology, buildings, trees	Rights Reserved - Free Access
Q4 2015	Národní památkový ústav	12,577 update (number of new items tbc)	Images, text	Archaeology	Rights Reserved - Free Access
Q4 2015	National Institute of Archaeology and Museum	800, update to enable display of thumbnails + enrichment	Images, text	Archaeology objects	Rights Reserved - Free Access
Q4 2015	Cultural and Educational Technology Institute - Research Centre Athena	3,900 update to fix display of thumbnails and to de-duplicate.	Images, 3D	Archaeology objects	CC-BY-ND
Q4 2015	Kulturstyelsen	301,314 Update to address broken links and enrichment (no new content at this time)	Images, text	Archaeology monuments	Public domain
Q1 2016	German Archaeological institute	183,687, Update to address broken links	Images, text	Archaeology monuments and objects	CC-BY-SA
Q1 2016	Direction des Monuments et des Sites de la Région de Bruxelles-Capitale	14,590 update (10,000 new)	Images, text	Archaeology, buildings, trees	Rights Reserved - Free Access
Q1-2 2016	National Heritage Board, Poland	Update to add c. 7,000 records	Images, text, 3D	Archaeology, monuments	CC-BY-NC-ND (178) CC-BY-SA (59)
Q1 2016	CIMEC, National Archaeological Record of Romania	4,050, new collection	Images, text	Archaeology, monuments	
Q1 2016	CIMEC, Listed monuments of Romania	5,093, new collection	Images, text	Archaeology, monuments	
Q1 2016	STARC	ca. 400, update to de-duplicate	Images, text, 3D	Archaeology, monuments	CC-BY-NC-ND
Q1 2016	Pamiatkovy urad Slovenskeij republiky	is aiming to update through Slovakiana			
Q1-2 2016	Swedish National Heritage Board	81,2971 + update to enhance spatial indexing and add new content (number tbc).	Text, images	Archaeology, Monuments	Public domain
	Total	**40,062**			

CARARE are in contact with other members of their network who may also wish to update their collections during the second half of the project.

2.5 EuropeanaFashion

Table 7: Planned ingestion for EuropeanaFashion

Month/Quarter	Data provider	Estimated number of records	Type	Subject	IPR (content)
Q3 2015	Victoria & Albert Museum	382	IMAGE	Fashion & costume objects	CC-BY-SA
Q4 2015	Amsterdam Museum	1,200	IMAGE	Fashion & costume objects, fashion prints	CC-0, public domain, Rights Reserved - Free Access
Q4 2015	Gemeentemuseum Den Haag	1,300	IMAGE	Fashion prints	Rights Reserved - Free Access
Q4 2015	TextielMuseum Tilburg	500	IMAGE	Fashion & costume objects	Rights Reserved - Free Access
Q4 2015	Modemuseum Hasselt	290	IMAGE	Fashion & costume objects	Rights Reserved - Free Access
Q4 2015	ITS International Talent Support	175	IMAGE	Catwalk photos, photography of fashion objects	Rights Reserved - Free Access
Q1 2016	Catwalk Pictures	40,000	IMAGE	Catwalk photos	Rights Reserved - Free Access
Q1 2016	Missoni	5,000	IMAGE	Catwalk photos, photography of fashion objects	Rights Reserved - Free Access
May 2016	MoMu	1,500	IMAGE	Fashion and Costume objects	Rights Reserved - Free Access
May 2016	Koninklijke Bibliotheek van België	2,000	IMAGE	Prints related to fashion	Rights Reserved - Free Access
May 2016	Museum Applied Arts Belgrado	300	IMAGE	Catalogues, Exhibitions, museum objects	Rights Reserved - Free Access
	Total	**52,647**			

2.6 EUscreen

EUscreen aggregation for DSI1 will commence from Q4 2015. There are two strands of work:
- new data providers. EUscreen is in touch with the data providers below, and is working on getting them familiar with the tools used by EUscreen. An online course was created to support this activity. Next to the online tutorial, a course is organised at the BUFVC premises (1st week of February 2016), where all the expected new providers, as listed below, are invited to. At this point, more insights are given in the exact dimensions of the contributions; notably the number, time and IPR status of the collections they will make available through Europeana.
- existing data providers. EUscreen will continue to work with the existing content partners and invite them to contribute to the existing body of content.

Table 8: Planned ingestion for EUscreen

Month/Quarter	Data provider	Estimated number of records	Type	Subject	IPR (content)
Q2 2016	Albanian State's Film Archives (AQSHF)	300	Moving image	Documentary & newsreels	Rights Reserved - Free Access
Q2 2016	Associated Press	550,000	Moving image	Newsreels	Rights Reserved - Free Access
Q2 2016	Centre National de l'Audiovisuel (CNA)	400	Moving image	Broadcast materials from the	Rights Reserved - Free Access

				region	
Q2 2016	European Broadcasting Union	TBD	Moving image	Eurovision Song Festival	Rights Reserved - Free Access
Q2 2016	Hrvatska Radiotelevizija Croatia (HRT)	735	Moving image and contextual items	Broadcast materials from the region	Rights Reserved - Free Access
Q2 2016	Maltese Broadcasting Authority	39	Moving image	Broadcast materials from the region	Rights Reserved - Free Access
Q2 2016	Moldova Film	TBD	TBD	Broadcast materials from the region	Rights Reserved - Free Access
Q2 2016	NRK	TBD	TBD	Broadcast materials from the region	Rights Reserved - Free Access
Q2 2016	Radio Television Vojvodina	TBD	TBD	Broadcast materials from the region	Rights Reserved - Free Access
Q2 2016	RTVE / Filmoteca Española	4,012	Moving image	No-Do newsreel collection	Rights Reserved - Free Access
Q2 2016	TV Slovenia	TBD	TBD	Broadcast materials from the region	Rights Reserved - Free Access
Q2 2016	University of Ulster	TBD	TBD	Broadcast material from the 1970s onwards - w focus on Northern Ireland	TBD
Q2 2016	Vlaams Instituut voor Audiovisuele Archivering (VIAA)	360,000	Newspaper pages	WW1	Rights Reserved - Free Access
Q2 2016	Yorkshire Film Archive	TBD	TBD	Broadcast materials from the region	Rights Reserved - Free Access
	Total	**915,486**			

2.7 HOPE2.0 / International Association of Labour History Institutions

Table 9: Planned ingestion for HOPE2.0

Month/Quarter	Data provider	Estimated number of records	Type	Subject
March 2016	Arbeiderbevegelsens arkiv og bibliotek (Arbark) (Oslo)	20,000	Images	Norwegian Labour Movement
March 2016	Arbeiderbevegelsens arkiv og bibliotek (Arbark) (Oslo)	500	PDF files	Minutes of Norwegian Congress of Trade Labour Unions. Reports from the Norwegian Confederation of Trade Labour Unions. Minutes of Norwegian Labour Party Rallies. Reports from the Norwegian Labour Party.
March 2016	European Trade Union Institute (ETUI) (Brussels)	2,700	PDF files	Documents concerning and publications by European Trade Unions and 'le monde social' in Europe.
March 2016	Archiv für Soziale Bewegungen (Bochum)	2,650	Images	Trade Unions and social movements, predominantly from the German Ruhr region, including the German Miners' Union, the German Chemical Workers' Union, etc.
March 2016	Trade Union Congress Library Collections at the London Metropolitan University (TUC) (London)	5,000	Images	Pamphlets and other material collected from unions, pressure groups and campaign movements both in the UK and internationally since the second half of the 19th century.

	Αρχεία Σύγχρονης Κοινωνικής Ιστορίας (Contemporary Social History Archives) (ASKI) (Athens)			
March 2016		24,185	TEXT	Culture and the Greek Left (1900-1974). The Greek Civil War (1946-1949). Greek Youth Movements (1936-1974). Greek Underground Press (1936-1974).
	Total	**55,035**		

2.8 The European Library

Table 10: Planned ingestion for The European Library

Month/Quarter	Data provider	Estimated number of records	Type	Subject
Q4 2015	"Lucian Blaga" Central University Library of Cluj-Napoca	583	TEXT	Old and rare books
Q4 2015	"Lucian Blaga" Central University Library of Cluj-Napoca	109	TEXT	Manuscripts
Q4 2015	"Lucian Blaga" Central University Library of Cluj-Napoca	40,056	TEXT	Periodicals
Q4 2015	"Lucian Blaga" Central University Library of Cluj-Napoca	1,057	IMAGE	Iconographic materials
Q4 2015	"Lucian Blaga" Central University Library of Cluj-Napoca	32	TEXT	University Library Publications
Q4 2015	"Lucian Blaga" Central University Library of Cluj-Napoca	396	TEXT	Books and reprints
Q4 2015	Croatian Academy of Sciences and Arts	260	IMAGE	Photography collection of the Fine Arts Archives
Q4 2015	Croatian Academy of Sciences and Arts	53	IMAGE	Photographic archive
Q4 2015	Croatian Academy of Sciences and Arts	2,645	TEXT	Academy Publications
Q4 2015	Croatian Academy of Sciences and Arts	19	IMAGE	Catalogs of the Strossmayer Gallery of Old Masters
Q4 2015	National Library of Czech Republic	46,085	TEXT	Manuscriptorium
Q4 2015	National Library of Czech Republic	796	TEXT	Kramerius - 19th and 20th century digitised periodicals and monographs from the Czech Republic
Q4 2015	National Library of Denmark	10,645	TEXT	Archive of Danish Literature
Q4 2015	National Library of Denmark	33,093	TEXT	Articles from the Tidsskrift.dk
Q4 2015	National Library of Russia	632	TEXT	Russian Book Treasures
Q4 2015	National Library of Russia	4,655	TEXT	Russian classics
Q4 2015	National Library of Russia	8,005	IMAGE	Digital History of Russia
Q4 2015	National Library of Russia	8,304	IMAGE	Imprint Collection
Q4 2015	National Library of Russia	3,235	IMAGE	Maps
Q4 2015	National Library of Russia	4,519	IMAGE	Diocesan Gazette
Q4 2015	National Library of Russia	2,808	IMAGE	Saint-Petersburg

Q4 2015	National Library of Russia	81	IMAGE	History of the Imperial Public Library
Q4 2015	National Library of Russia	34	IMAGE	Children's books in Yiddish
Q4 2015	National Library of Scotland	5,000	IMAGE	Historic map collection
Q4 2015	National Library of Scotland	380	TEXT	Scottish Post Office directories
Q4 2015	National Library of Scotland	80,000	IMAGE	Modern maps
Q4 2015	National Library of Scotland	1,700	TEXT	Early Gaelic Books
Q4 2015	National Library of Scotland	200	TEXT	British military lists
Q4 2015	National Library of Scotland	1,770	TEXT	Word on the street broadside collection
Q4 2015	National Library of Scotland	200	TEXT	Books
Q4 2015	National Library of Scotland	350	IMAGE	Albums of the Edinburgh Calotype Club
Q4 2015	National Library of Scotland	100	TEXT	Medical history of British India
Q4 2015	National Library of Scotland	40	IMAGE	The Forth Bridge
Q4 2015	National Library of Scotland	91	IMAGE	The Tay Bridge
Q4 2015	National Library of Scotland	696	IMAGE	Jacobite ballads
Q4 2015	National Library of Scotland	2,432	IMAGE	English ballads
Q4 2015	National Library of Scotland	156	IMAGE	Edinburgh theatre posters
Q4 2015	National Library of Scotland	34	IMAGE	Edinburgh theatre photos
Q4 2015	National Library of Scotland	537	IMAGE	Hutton drawings
Q4 2015	National and University Library in Zagreb	3,485	TEXT	NLH:Europeana Newspaper
Q4 2015	National Library of Spain	20,000	TEXT; IMAGE; (SOUND)	Biblioteca Digital Hispanica
Q4 2015	National Library of Spain	30,000	TEXT	Digital Periodicals Library
Q4 2015	Royal Library of Belgium	30,000	TEXT	KBR: Europeana Newspaper
Q4 2015	National Library of Turkey	100,000	TEXT	NLT: Europeana Newspaper
Q4 2015	Russian State Library	12,000	IMAGE	Sheet music
Q4 2015	National Library of Latvia	20	IMAGE	Glass plate photographic negatives
Q4 2015	National Library of Bulgaria	11	TEXT	Digital Library of Bulgaria - Slavonic manuscripts
Q4 2015	National and University Library of Slovenia	58,023	TEXT	Laibacher Zeitung
Q4 2015	Swiss National Library	5,500	TEXT	Digitized publications from e-Helvetica
Q4 2015	The University and National Library of Debrecen	11,832	TEXT	18th and early 19th Century Hungarian Periodicals
Q4 2015	The University and National Library of Debrecen	12,563	TEXT	Playbills of the Debrecen Theatre 1862-1920
Q4 2015	The University and National Library of Debrecen	4,551	TEXT	19th Century Periodicals from the Eastern Hungarian Region
Q4 2015	Tilburg University	1,658	TEXT/ DATASET	Dataverse

	Total	551,431		

3. Ingestion for EU funded projects

The following table lists the running EU funded projects with their start and end dates. The total number of expected records are listed together with the current amount of records in the Europeana portal.

Table 11: Total planned ingestion for EU funded project and status quo as of July 2015

Project name	Start	End	Total records (DoW)	In portal (July 2015)
Europeana Cloud*	1-2-2013	1-2-2016	2,400,000	1,540,653
EUscreenXL	1-3-2013	29-2-2016	1,000,000	543,089
LoCloud	1-3-2013	29-2-2016	4,000,000	244,924
EAGLE	1-4-2013	31-3-2016	1,500,000	346,344
Europeana Food and Drink	1-1-2014	31-8-2016	50,000 - 70,000	0
Europeana Sounds**	1-2-2014	31-1-2017	770,000	30,255

* Europeana Cloud delivers via TEL (see previous chapter).
** Europeana Sounds will deliver 545,000 sound items and 225,000 supporting items.

Table 12: Amount of metadata records to be provided between October 2015 and the end of the project with an average per month. The average is calculated as number to be provided devided by number of remaining months. Actual submission per month may be different.

Project name	Still to provide	Average per month
Europeana Cloud	859,347	215,000
EUscreenXL	456,911	92,000
LoCloud	3,755,076	750,000
EAGLE	1,153,656	192,000
Europeana Food and Drink	50,000 - 70,000	4,500 - 6,300
Europeana Sounds	739,745	46,000

4. Ingestion for other aggregators (national and thematic) and direct providers

Europeana launched an aggregator survey to map the aggregation landscape for digital cultural content across Europe, both to provide an overview of the current situation and to inform future strategic planning to promote resilience, sustainability and innovation. This survey was open for four weeks between 8 January and 6 February 2015. It was completed by 33 aggregators, which together provide more than 80% of the data that is currently available in Europeana. The complete summary of the survey results is provided in D1.1 of Europeana version 3[7].

Almost two thirds (60%) of the aggregators have an ingestion plan, in which they schedule the amount of data to be processed. Most of the respondents did, nevertheless, answer the question of how many digital objects they plan to deliver to Europeana in 2015. The majority plans to make a substantial contribution this year: between 10,000 and over one million objects. The below table is a detailed overview of the number of records the partners planned to deliver in early 2015 with an overview of how much they have submitted during the year.

Table 13: Expected ingestion for other aggregators

	Planned records	New in portal*	Frequency
RNOD	100,000 - 500,000	0	1 - 5 times a year
Arts Council Norway	> 1,000,000	532,129	Every month
Digitale Collectie	> 1,000,000	903,358	Every month
HISPANA	100,000 - 500,000	129,269	6 - 11 times a year
Swedish Open Cultural Heritage / K-samsök	> 1,000,000	233,985	Every month
Deutsche Digitale Bibliothek	> 1,000,000	10,025	
Slovenski nacionalni agregator e-vsebin	10,000 - 100,000	64,067	1 - 5 times a year
Hungarian National Digital Archive and Film Institute (MandA)	10,000 - 100,000	20,501	Every month
Cyprus Ministry of Education and Culture (CMOEC)***	100,000 - 500,000	4,602	1 - 5 times a year
Hellenic Aggregator for Europeana	1,000 - 10,000	16,572	1 - 5 times a year
Danish National Aggregator	100,000 - 500,000	16,112	1 - 5 times a year
Public Library - Varna	10,000 - 100,000	22,820	1 - 5 times a year
CULTURAITALIA**	500,000 - 1,000,000	? (via AthenaPlus)	1 - 5 times a year
Bulgariana	100 - 1,000	0	1 - 5 times a year
Erfgoedplus.be	1,000 - 10,000	5,496	Every month
Národné Osvetové Centrum (NOCKA) (not active yet)	1,000,000	0	1 - 5 times a year
EuropeanaLocal Austria	10,000 - 100,000	1,055	6 -11 times a year
Rijksmuseum	20,000	0	
Total	**~6,000,000 - 7,500,000**	**1,959,991**	

* January - August 2015
** CULTURAITALIA delivers via AthenaPlus. At the time of writing it was not possible to see the number of records that was submitted in early 2015.
*** CMOEC is currently preparing a larger submission of data for Europeana.

---

[7] D1.1: Recommendation to Improve Aggregation Infrastructure:
http://pro.europeana.eu/files/Europeana_Professional/Projects/Project_list/Europeana_Version3/Deliverables/EV3%20D1_1%20Aggregation%20Infrastructure.pdf

Conclusion and recommendations

The table below shows a summary of totals of expected metadata records from October 2015 until June 2016. In total almost 13 million records are to be expected; this would be an average of 1.4 million records per month, which is quite high if you compare it to the growth curve of 908,178 records per month in during the first half year of 2015 (see paragraph 1.5). As there are still a lot of gaps in the planning this can be seen as a rough estimate. It is recommended to compare actual numbers in July 2016 with the numbers in this ingestion plan to check its accuracy.

Table 14: Estimation of expected metadata records for all aggregators

Year	Quarter	DSI quarter	Month	Expected number of records	
2015	Q4	Q2	October	359 (DSI)	588,903 (DSI)
			November		3,898,500 (projects)
		Q3	December	5,650 (DSI)	4,040,009 (other)
2016	Q1		Januari		509,808 (DSI)
			Februari		2,626,500 (projects)
		Q4	March	884 (DSI)	
	Q2		April		978,770 (DSI)
			May		151,500 (projects)
			June		
Total					**12,793,990**

Having in mind that we need to focus a lot more on the quality of the data, the monthly growth rates indicated above are worrying. As we cannot handle big quantities of data in a meaningful way, we need to find ways moving the focus of our work to quality data. A content strategy will be developed in the next six months that will also inform the ingestion planning for the future. A Data Quality Committee recruited from members of the Aggregator Forum will support the content strategy development with specific quality indicators and targets to both improve the quality of our legacy data and new submissions from all our partners.

Again it proved to be a difficult exercise to create an ingestion plan due to several reasons. For non-project based aggregators, a long-term planning is very difficult as they are also dependant on their data partners. The uncertainties of the Europeana ingestion plan reflects the uncertainties of the aggregators ingestion plans and is a consequence of the current aggregator model. For projects the number of metadata records has been decided upon in the Description of Work, but it is not always possible to make a detailed planning of when and how often metadata will be delivered.
Another issue for making an accurate ingestion plan is the different timeline aggregators are working towards. Projects have clear deadlines and data submission increases towards the end of the project. Projects as well as DSI partners do not plan in calendar years but in project months. This makes it difficult to align a plan on a quarterly basis.

As long-term ingestion planning is very difficult, it is recommended to find a more time sensitive way of planning. This could be in the form of a living document which can be updated several times a year or by making multiple short-term (e.g. quarterly) ingestion plans. The European Library is working in a similar way and we will explore how we can benefit from the experience of The European Library. We are also exploring the possibility of setting up a full ingestion backlog to keep track of submissions but also keep track of plans for data submissions in a dynamic way. Now that our ingestion infrastructure allows us to publish data in a much more flexible way, we also need to find a more flexible way to plan our ingestion activities, which a backlog could help with. Instead of an ingestion plan like the current document, which has a strong information component, an ingestion backlog would become a real practical tool to plan ingestion work much better. We are working on this together with our data partners in the coming months and share our experiences with our next ingestion plan milestone.