

ECP-2007-DILI-517005

ATHENA

Final Report on networking framework of non-partner projects and other bodies

Deliverable number	<i>D5.4</i>
Dissemination level	<i>Public</i>
Delivery date	<i>30 April 2011</i>
Status	<i>Final</i>
Author(s)	<i>Marzia Piccininno (MiBAC) Maria Teresa Natale (MiBAC) Valentina Vassallo (The Cyprus Institute) Annette Friberg (Europeana)</i>



eContentplus

This project is funded under the eContentplus programme¹,
a multiannual Community programme to make digital content in Europe more accessible, usable and exploitable.

¹ OJ L 79, 24.3.2005, p. 1.

Table of Contents

FOREWORD	3
OBJECTIVE OF THE DELIVERABLE	3
1. INTRODUCTION	5
1.1 OVERVIEW OF THE SURVEY	5
1.2 METHODOLOGY	6
2. ANALYSIS	7
2.1 GENERAL INFORMATION	7
2.2 CONTENT INGESTION	13
2.3 PARTNERS RELATION	15
2.4 IDENTIFIERS	16
2.5 MULTILINGUALISM	16
2.6 END-USER SERVICES	18
2.7 FINANCE AND SUSTAINABILITY	20
2.8 OTHER	21
3. CONCLUSIONS	22
ANNEX I: THE QUESTIONNAIRE	24
ANNEX II: AGGREGATORS' LIST	33
ANNEX III: POSITIONING	36
ANNEX IV: CONTENT PROVIDERS OF EACH AGGREGATOR	38
ANNEX V: PLANNED DIGITAL ITEMS TO BE DELIVERED TO EUROPEANA .	40

Foreword

This report is a result of a joint work carried out by ATHENA and Europeana. Both projects have interest in investigating the world of the European aggregators: Europeana aims at working with huge aggregators in order to retrieve all at once critical masses of normalised data coming from many memory institutions, otherwise the interactions with the thousands European cultural bodies, small and large, national and local, would be impossible. Of course, Europeana knows very well the aggregators funded within the European Commission programmes whose main goal is supply content to the portal; however, many other aggregators exist across Europe (thematic, national cross-domain, regional, etc.) and from the time of the previous similar survey (end of 2009) many others were born.

By the other hand, ATHENA provided Europeana many content coming from various aggregators (domain aggregators specific for museum objects and national aggregators, the so called culture portals) and has acquired specific knowledge on how they are structured; furthermore, the evaluation of projects with similar and comparable aims facilitates the content aggregation process, avoids overlapping and promotes the circulation of best practices among the Europeana project family.

Objective of the deliverable

The ATHENA deliverable 5.2 ‘Guidelines and best practices setting up National Cooperation Frameworks (NCFs)’ (October 2009) highlighted how the existence of NCFs across Europe is essential for the sustainability of Europeana, and how they are mainly realised through national thematic or cross-domain aggregators that implement effective cooperation strategies among memory institutions.

By the other hand, the ATHENA Description of Work states that among the activities to be carried on in relation to the implementation of Europeana, there are:

- working in joint cooperation with the other ongoing trans-European and national projects, to fully support the development of Europeana;
- surveying and look for the most advanced available technologies suitable for museums, and experiment their integration into concrete services;
- increasing interconnection and interoperation of national cultural networks.

In fact, the ATHENA WP5 activity is a continuous effort for the harmonisation of the activities run in ATHENA with the specific needs of the content providers and, beyond, for integrating the results achieved by the project into the scheme of Europeana. All efforts are realised taking maximum advantage from cooperation with other projects working with similar and comparable aims.

For all these reasons it seemed appropriate to include as annex the Europeana and ATHENA Survey for Aggregators #1 that was separately concluded at that time and that had quite a stir by the Europeana office because it represented a guide to the world of aggregators that stated interacting with the European portal.

The information gathered within the Europeana and ATHENA Survey for Aggregators #1 in 2009 was really helpful for ATHENA too for establishing cooperation channels with some

projects aiming at the Europeana content provision. Knowing each other and disseminating outcomes and tools helped into the aggregation process. For instance, now the MIMO project uses LIDO, the new format for data harvesting and interchange for museums digital objects developed within ATHENA¹, or Judaica Europeana employs the ATHENA ingestion software to upload the museum content and the APENet tools for the archive ones, or more, ATHENA, thanks to APENet, found a better way to supply Europeana with the national archive contents².

The outcome of the Survey for Aggregators #1 went beyond the expectations because it was a substantial background for structuring the Europeana content development strategy and helped the ATHENA consortium to have a wider view on the other European networks and understand their operational mechanisms.

For this reason it appeared appropriate to arrange for deliverable 5.4 'Final report on networking framework of non partners projects and other bodies' a second Europeana and ATHENA Survey for Aggregators to be released at the end of the project (M30) as heritage for future cooperation developments and sustainability strategies³.

¹ www.lido-schema.org/

² See deliverable 1.3 'Second Periodic Report' (April 2010), p. 10 and deliverable 1.5 'Final report'.

³ The analysis carried out in this deliverable could be considered as an integration of deliverable 1.5 'Sustainability concept'.

1. Introduction

At the end of 2009 the ATHENA and Europeana projects published a report on the state of the art of the European aggregators (published by ATHENA as annex to deliverable 5.2 ‘Guidelines and best practices setting up National Cooperation Frameworks (NCFs)’, October 2009 and included as reference in the European Aggregators’ Handbook (May 2010)¹.

The original idea of sharing good practices in content aggregation emerged during an informal meeting that was held in March 2007 by the Institut für Museumsforschung-SPK in Berlin and had a formalisation during the aggregator workshop organized in Rome on June 16th, 2009 by SPK and the Ministry for Cultural Heritage and Activities-ICCU (Italy), coordinator of the ATHENA project were many European managers and technicians dealing with national cross-domain portals or digital libraries discussed their functionalities in comparison to Europeana and exchanged experiences².

This meeting showed that the cultural institutions involved in aggregation of content – including Europeana - share similar approaches and technical solutions. As a consequence, it seemed opportune to further investigate this matter with the Europeana and ATHENA Survey for Aggregators #1.

The analysed results were discussed during the Europeana Plenary Conference in September 2009 and the Aggregators

Round Table organised in Lund (Sweden) one month later³.

In this light, it is opportune to recall the aggregator definition as given by Europeana in the Aggregators Handbook: “An aggregator in the context of Europeana is an organisation that collects metadata from a group of content providers and transmits them to Europeana. Aggregators gather material from individual organisations, standardise the file formats and metadata, and channel the latter into Europeana according to the Europeana guidelines and procedures. Aggregators also support the content providers with administration, operations and training”.

From 2007 onward many things changed: Europeana was launched in November 2008 and many other aggregators, in particular national cross-domain aggregators, were born. For this reason it seemed opportune both to Europeana and ATHENA to repeat the survey in order to monitor the evolution of the aggregator landscape and to reinforce cooperation among them.

The Survey for Aggregators #2 was released at the end of 2010.

1.1 Overview of the survey

The questionnaire of the Europeana and ATHENA Survey for aggregators #2 was elaborated on the basis of the old one; the greater part of the answers was left as they were and the remaining ones were improved. It deals with the main areas of establishing a partnership between the various aggregators.

1 http://www.europeana-libraries.eu/c/document_library/get_file?uuid=94bcddb3-3625-4e6d-8135-c7375d6bbc62&groupId=10602.

2 Minutes and .pps presentations are available at the [URL](http://www.athenaeurope.org/index.php?en/111/events/61/rome-working-meeting-on-cross-domain-aggregators-in-europe) <http://www.athenaeurope.org/index.php?en/111/events/61/rome-working-meeting-on-cross-domain-aggregators-in-europe>.

3 Programme and presentations at <http://group.europeana.eu/web/europeana-project/roundtable>.

The survey consists of 77 questions divided into 8 parts (see the complete questionnaire in ANNEX I), focused more on workflow, aggregation strategies and service policies rather than technicalities:

1. General information
2. Content Ingestion
3. Partner Relations
4. Identifiers
5. Multilingualism
6. End-User Services
7. Finance and sustainability
8. Others (relation with Europeana)

The Europeana and ATHENA projects launched the survey with the purpose to verify the shared issues and enable the establishment of aggregators in contributing content to Europeana. The Survey wants also to investigate possible relations, similarities and parallelisms with ATHENA and other projects aimed at content provision.

1.2 Methodology

The questionnaire distributed across the end of 2010 and the beginning of 2011 to known aggregators of digital cultural heritage content. The addressees were the aggregators that answered the first questionnaire and the other brand new ones that were born in the meantime across the European Union. Institutions had to compile the questionnaire online¹.

The person that had to supply the answers was identified as the person in charge of strategy or IT/digitisation in the organization, supported – when necessary – by metadata and technical advisors.

30 aggregators answered the questionnaire in 2009; 23 of them shared this second round too. 43 total feedbacks were received for the Survey 2011 (see the complete list in ANNEX I I).

The results of this research suggest the strategies, activities and services that need to be developed to help aggregators act as portals in their own right and help in the production of Europeana.

¹ <http://www.surveymonkey.com/s/W9KBZVW>

2. Analysis

This report will give an overview of the outcomes, section by section.

Quantity information is often given between brackets, e.g. ‘Many aggregators provide digital object storage too for their content providers (22)’ means that 22 aggregators provide storage services.

Please note that **the respondents skipped some questions**.

2.1 General information

43 answers were collected. All the initiatives are **based in European countries** (see ANNEX II):

1. Archiefbank
2. BAM
3. BHL-Europe
4. Carmentis
5. CER.ES
6. Collections
7. CulturalItalia
8. Culture Grid
9. Digital Libraries of Serbia
10. Digital Library of Slovenia
11. DISMARC
12. ECLAP
13. EFG - The European Film Gateway
14. Encycloasturias - Mediateca de la Universidad de Oviedo
15. Erfgoedplus.be
16. eSbirky
17. EuropeanaLocal Austria
18. Expo @ AcrossLimits
19. Gateway Bayern
20. Hellenic Aggregator for Europeana
21. Hispana
22. HOPE project
23. Judaica Europeana
24. Kulturav Västernorrland
25. Kulturpool
26. Kypriana
27. Letonica
28. Manuscriptorium

29. MIMO (Musical Instruments Museum Online)
30. MovE
31. Muis
32. National Digital Library of Finland
33. PIONIER Digital Libraries Federation
34. Porstua Web Service
35. Portail de la musique contemporaine
36. Public Library - Varna
37. RNOD - Registo Nacional de Objectos Digitais
38. SCRAN
39. SOCH
40. The European Library
41. Thuis in Brabant
42. Vlaamse Kunstcollectie
43. www.africamuseum.be

Geographical level of aggregation

The geographical level of aggregation was intended as the level at which the organisation operates in terms of the data it collects i.e. whether the sources are regional, European or worldwide (FIG.1). 22 aggregators collect content at national level but only 16 are the so-called “national portals” i.e. national cross-domain aggregators working across several sectors and collecting material from museums, libraries and archives.

GEOGRAPHICAL LEVEL OF AGGREGATION	# OF AGGREGATORS
National	22
Worldwide (EU countries and beyond)	10
Regional	9
European (only EU countries)	2

FIG. 1 - Geographical level of aggregation

In comparison with the survey 2009 (FIG. 2), the collected answers outline a little bit different geographical level of aggregation: a considerable amount of aggregators (10) collect content at worldwide level (but are based in Europe), and the regional aggregators that were underrepresented in Survey #1 now tripled and make the statistical information more consistent with the European reality.

GEOGRAPHICAL LEVEL OF AGGREGATION	% 2011	% 2009
National	51	60
Worldwide (EU countries and beyond)	23	13
Regional	21	7
European (only EU countries)	5	20

TABLE 2 - Geographical level of aggregation: comparison results 2009 and 2011.

Type of aggregation

The greater part of the aggregators (23 – 51%) aggregates cross-domain content i.e. it works across several sectors and collects material from museums, libraries and archives.

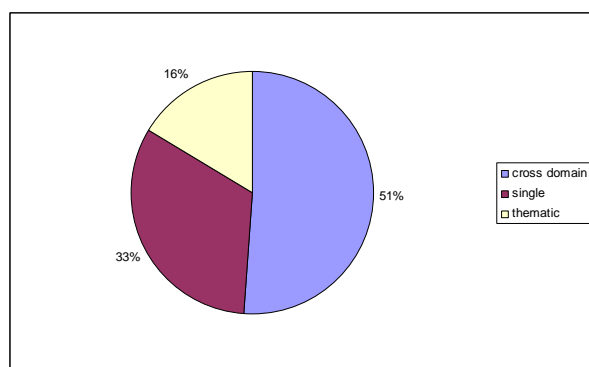


FIG. 3 – Type of aggregation

14 aggregators (32.5%) are single domain aggregators because they represent and collect data only from one cultural field (FIG. 3). Contrary to the expectations, the single domain aggregators don't represent only the library sector, which is the most known, but museums too (FIG. 4).

Single domain aggregators	
Library	6
Museum	6
Archive	1

FIG. 4 – Single domain aggregators by sectors

7 aggregators are focused on specific themes: biodiversity, performing arts, social and labour history, manuscripts, music instruments, the Jewish participation in the development of urban culture in Europe, and contemporary music. Obviously many of them are cross-domain aggregators, however in those cases the theme is prevailing on the kind of content.

Online availability of content

The largest part of the aggregators (34 = 79%;53% within survey #1) has already a public interface and their content can be browsed online, while 4 of them will be launched during 2011 (BHL-Europe, EFG, the National Digital Library of Finland, and RNOD).

Only 4 of them don't foresee this option of being online because they act as a bridge towards other portals or applications: the Hellenic aggregator, the HOPE project and MIMO work only for Europeana, while SOCH, the system of the Swedish Heritage National Board, doesn't have its own portal because the content comes from some 40 institutions and can be accessed via a the SOCH API and data is so far used in 8 different applications.

Content in Europeana

30 aggregators out of 43 have already delivered their content to Europeana: 17 did it thanks to a direct link to Europeana, 9 through other projects or aggregators, while 4 chose both ways (directly to Europeana and through other aggregation initiatives). It is the case, for instance, of CulturalItalia that sent their metadata through OAI-PMH protocol to Europeana directly as well as through the ATHENA project.

Some projects or aggregators aggregated and supplied Europeana with content coming from other aggregators:

The European Library delivered content of

- Digital Library of Serbia
- Digital Library of Slovenia
- Letonica
- Manuscriptorium

ATHENA

- BAM
- CulturalItalia
- eSbirky

Hispana

- Cer.es

Europeana Connect

- DISMARC

EuropeanaLocal

- Ergoedplus
- MovE
- Public Library – Varna

BHL

- Africamuseum.be

MIMO

- Africamuseum.be

STERNA

- Africamuseum.be

This datum is particularly important because it demonstrates that the ‘virtuous circle of aggregation’ for circulation and reuse of cultural information that Europeana wished has started up.

Amount and typology of aggregated digital items

Question 9 of the survey asked aggregators how many digital items they aggregated at the time of the questionnaire¹ (FIG. 5).

The answers showed a very uneven amount of aggregated metadata, from 100 of Encycloasturias to over 4 million of the French portal Collections.

¹ Europeana counts the number of records. So a document with 300 pages through which you can browse on the local site but which has metadata only at the level of the document will count as one digital object and one record in Europeana. Pages from a book that have separate meaningful metadata useful for finding the individual page may be counted as separate digital objects and records.

FIG. 5 – Aggregated items and expected increase 2011

AGGREGATOR	AMOUNT OF AGGREGATED ITEMS	EXPECTED INCREASE IN 2011
Archiefbank	7,677	-
BAM Portal	4,000	+ 7,320 (4 collections)
BHL – Europe	89,860	Total to be aggregated by April 2012: 109,618 (18 collections)
Carmetis	6,000	41,700 (7 collections)
CER.ES	105,537	around 30,000
Collections	4,252,537 documents 2,907,036 images	-
CulturaItalia	950,000	500,000 (from Regions, private archives, Institutes of the Ministry)
Culture Grid	546,964 (aggregated to Europeana); 1,126,254 (item records in Culture Grid)	30,864 (4 collections)
Digital Libraries of Serbia	50,000	20,000 - from other regional libraries in Serbia
Digital Library of Slovenia	450,000	3 portals are planned
DISMARC	100,000	-
ECLAP	-	1,000,000 items

EFG European Film Gateway	30,000	Total to be aggregated by August 2011: 280,875 (15 collections)
Encycloausturias	100	2,300
Erfgoedplus.be	50,000	estimation 25,000 from many sources
eSbírky	30,813	500,000 from Czech museums
EuropeanaLocal	10,000	60,000 (from various universities, local museums, archives, libraries)
Expo @ Acrosslimits	50,000	100,000 (Private Collectors, Private Museums, Photographers, Other individuals, General Public)
Gateway Bayern	350,000	80,000 from libraries
Hellenic Aggregator for Europeana	128,000	Unknown
Hispana	1,234,252	500,000 (21 collections)
HOPE project	Not applicable yet; estimate is nearly 3 million files for the project	0

Judaica Europeana	6,559	225,965 describing 5 million information about
Kulturarv Västernorrland	62,000	28,500 (5 collections)
Kulturpool	250,000	173,000
Kypriana	3,000	-
Letonica	1,864	25,270 -
Manuscriptorium	37,762	about 40,000 - It depends on funding
MIMO-DB	21,135	47,100
MovE	130,000	About 10% (all partner-museums)
Muis	1,446,034	4,000 (via ATHENA)
National Digital Library of Finland	-	-
PIONIER Network Digital Libraries Federation	480,000	175,000 to 200,000
Porstua Web Service	4,000	ca. 1,000
Portail de la musique contemporaine	35,807	1,000
Public Library – Varna	12,000	-
Registo Nacional de Objectos Digitais – RNOD	Not yet applicable	Not yet applicable
Scran	360,000	Project dependant

SOCH	1,200,000	Via Carare: 880,000 records of historical buildings and ancient monument records.
The European Library	2,000,000	500,000 digital objects from Research Libraries via the Europeana Libraries project
Thuis in Brabant	110,000	250,000
Vlaamse Kunstcollectie	8,000	1,000
www.africanmuseum.be	For BHL Europe we have currently 150 books (+/- 50,000 pages). For STERNA information on some 100 bird specimens and 14 ethnographic objects (prototype project on using semantic Web)	In BHL-Europe target is some additional 50,000 pages (depends on national funding). In Open-UPs (new project) target is 15,700 photographs of natural sciences, 1,000 drawings, about 50 audio files, and 700 video files, 300,000 aerial photographs

The kind of aggregated items considered for this survey are the four that Europeana

considers: audios, videos, texts, and images.

26 aggregators give access to various kinds of content, although only 2 (DISMARC and EuropeanaLocal) contains at the same time metadata of audios, videos, texts and image.

Metadata of images and texts are the most common aggregated items. 9 aggregators give access to 100% images, and 3 to texts only; in addition, 12 aggregators contain images predominantly (from 80 to 99%) and 8 texts. Metadata of audios and videos are rarely predominant: DISMARC contains 95% audios, while Encycloasturias 90% videos.

The percentage of other aggregated items (3D, catalogue records, etc.) is minimal.

Year of launch

Beside the veteran SCRAN (1996), the other aggregators have been launched from 2003 onward with a peak value between 2008 and 2011 when 28 aggregators started up. Many aggregators that were under construction on the occasion of Survey #1 came out (e.g. Expo@Acrosslimits).

Storage

Many aggregators (28) provide digital object storage too for their content providers if they can't maintain a repository; however this storage activity is not carried out extensively: 6 out 8 select the collections or the providers to store (e.g. CulturalItalia will do it in 2011 only for museums, Manuscriptorium only for the Czech National Library and not for other content providers); one aggregator, Vlaamse Kunstcollectie, outsourced the service to a private firm.

If aggregators don't store full digital content, they obviously provide external

links to the digital objects on the web sites of the content providers.

Other aggregators

Question 15 of the survey asked respondents to nominate other European aggregators they were aware of in order to verify other potential partners for Europeana. The majority of the feedbacks received showed that the overall part of the European aggregators is well known by other ones and Europeana itself. However some aggregators not yet involved into the Europeana galaxy came out (FIG. 6).

STATE	NAME	DOMAIN	URL
Italy	Archivi del Novecento	Archive	www.archividelnovecento.it
Germany	Deutsche Digitale Bibliothek	library	http://www.deutsche-digitale-bibliothek.de/
Sweden	Nationell ArkivDatabas	Archive	http://nad.ra.se/
European	Monasterium	archive	http://www.monasterium.net
Belgium	Amsab	archive	http://www.amsab.be/
France	BORA	archive	http://daf.archivesdefrance.culture.gouv.fr/sdx-222-daf-bora-ap/ap/index.xsp

FIG. 6 – Aggregator not yet involved into Europeana

Institutional support

The aggregators are supported by larger organisation of different nature, usually a

single institution although there are few exceptions (in particular EC funded projects) (FIG. 7).

Ministries, in particular ministries of culture, support 11 aggregators or digital libraries of national level (except Kypriana which is totally supported by a non-profit research institution). Ministries can also act in cooperation with other cultural institutions (4 cases).

Some aggregators are supported by a consortium funded within the European Commission programmes (6), other by museums only (5), libraries (4), heritage agencies (3), foundations or charities (3), national/local governments (2), and universities (1).

SUPPORTING ORGANISATION	# AGGR.
Ministry	11
European Commission	6
Museum	5
Library	4
Agency	3
Foundation/charity	3
Ministry + 1 other body	3
National/local government	2
EC + foundation/charity	1
Ministry + 3 other bodies	1
Museum + local government	1
University	1

FIG. 7 – Organisations supporting the aggregators

The general overview that emerged from this first section of the questionnaire is that cross-domain aggregators – national and regional - aim at improving online availability of the electronic information resources of libraries, archives and museums and developing functionalities to retrieve integrated information from all domains.

These action lines are usually supported by activities in aid of documentation/information, access/awareness, storage, and sometimes long-term preservation of the heritage.

Domain aggregators have a more specific goal such as provision of authoritative tools for documentation and discovery of specific items or topics (Jewish culture, social history, etc.).

Relation with Europeana

Question 17 asked aggregators how they position themselves in relation to Europeana: answers were often obvious (I'm a national aggregator, I'm a partner), but sometimes they expressed the pride to be small but important for the European portal. See the full list of answers in ANNEX III.

2.2 Content ingestion

Content strategy and ingestion plan

Only 19 aggregators programme their activities according to a content strategy or a collections development policy; 16 are developing it little by little that the works progresses, 6 don't have it, 2 skipped the question. However the good news is that 10 aggregators are used to publishing the strategy on their website.

22 aggregators have an ingestion plan with scheduling and amounts to keep under control the ingestion procedures.

Topics and categories of objects

Arts and humanities are the most covered areas of content (34 and 30 feedbacks respectively), social sciences and scientific areas follows immediately after (22 and 20 each).

Photos and pictures are the most common categories of aggregated material (88.1 and 85.7% each), followed by books (64.3%), audio files (57.1%), and journals (50%). 3D content is the less common (11.9%) (FIG. 8).

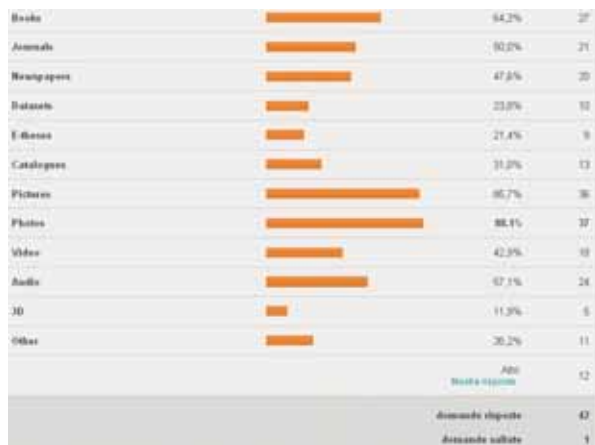


FIG. 8 – Categories of aggregated objects

The option ‘Other’ that the question foreseen gave some interesting results:

- play lists, animations, and interactive content;
- geodata for immobile heritage;
- moving images other than video;
- websites and virtual exhibitions;
- metadata of physical objects;
- online biographies and brief analyses of music works, concert ephemera (program notes);
- to all sorts of records / archival material;
- maps;
- documents related to film production, e.g. screenplays, dialogue lists, correspondence.

Metadata

Dublin Core is still the most used metadata schema for the object description of all cultural domains (26 aggregators use it), sometimes in association with other ones (LIDO, ESE, EAD, etc. - 12 cases).

8 aggregators use ESE¹ and other metadata schemas as format for the data exchange with Europeana; 7 did the same with LIDO², the harvesting format for museums developed within the ATHENA project.

It is worth to mention that 7 aggregators apply own metadata schemas, customised according to the specific domain requirements (for instance, European Film Gateway developed the EFG metadata schema composed of other standards - EN15907, ESE – in order to meet the necessity of the video archives).

METS, ISAD/ISAAR, MARC 21, MODS, ABCD, Darwin Core, ENRICH TEI P5, and PND5-DCAP are other occasionally applied formats.

Aggregation and export technical procedures

The OAI-PMH protocol is definitively the most common tool for the aggregation or harvesting of metadata (32), followed by FTP (13) and SRU (5). Some aggregators reported alternative ways to perform this task: in XML format (3), through the ATHENA ingester towards Europeana (2), Z39.50 (1), email or CD (1), REST/SOAP API (1), API and SWORD (1), and application developed with Liferay Webscraping (1).

Many aggregators declared that they are able to export their content for the reuse (22 cases); in addition, the only metadata export is even more common (34 cases)³. OAI-PMH is absolutely the most frequent way also to export content (18) and metadata (29) (FIG. 9).

¹ Carmentis, EuropeanaLocal, Expo, Hellenic aggregator for Europeana, Hispana, Kulturav Västernorrland, Kulturpool, the European Library.

² BAM, Carmentis, EuropeanaLocal, Kulturav Västernorrland, Kypriana, MIMO, National Digital Library of Finland.

³ EFG allows this only for Europeana.

Export Mechanisms	For content	For metadata
OAI-PMH	18	29
API	4	4
Opensearch	3	3
FTP	2	5
SRU	0	6
RSS FEED	-	5
Other	2	3

FIG. 9 – Export mechanisms

Software solutions

Open source software is largely used (25) but also proprietary software is common (22); only 3 aggregators based their systems on free software¹

¾ about of the aggregators (32) knows the existence of specific tools for the data ingestion like the ATHENA ingester and the Europeana content checker and 15 use them.

2.3 Partners relation

Institutions involved

The amount of cultural institutions that participate to the European aggregators as content providers is various, from the 1 of Gateway Bayern to 300 of PIONIER and SCRAN. A total of 1,934 cultural institutions of every kind and level are involved by the European aggregation initiatives (see details in ANNEX IV).

The aggregators usually actively approach new content providers (36); the criteria

followed for this enlargement procedure are the relevance of the collections to the aggregator enlargement policy above all (23), then the importance of the institution (22) or the collections' topics (16). The technical qualities of the content providers – development of technical functionalities (8) or continual supply of content (9) – are less considered criteria.

Licences and agreements

25 aggregators regulate the relationships with content providers through agreements but only 20 managed to provide a copy or an Internet reference.

The Creative Commons standard licences are the most used, in particular the CC by SA (Creative Commons by Share Alike) non Commercial (9), then CC by SA (2) and CC (1). Other kind of applied licences are customised to the specific institutional requirements (9) but sometimes modelled upon existing ones, the Europeana agreements for content providers and aggregators (3)².

When a specific licence is not applied, the relationships with the content providers are usually regulated thanks to general framework cooperation agreements or memorandums of understanding between the institution providing content and the body managing the aggregator (9); very few aggregators send formal letters or email when they need to ask the permission of use and reuse metadata (3), and 1 acts upon confidence!

When used, licences or other kind of agreements are mostly applied metadata, digital objects, and thumbnails (12), to

¹ The 3 software definitions of question #29 (open source, free and proprietary software) are those ones given by GNU (<http://www.gnu.org/philosophy/categories.html>).

² At the time of this report the new Europeana data exchange agreement (DEA) was under finalisation. The aggregators refer to the previous licences for content providers and for aggregators.

metadata only (7) or to metadata and thumbnails (5), depending on the way the aggregators are structured (some aggregate metadata only, some digital objects too). Rarely, licences are applied to digital objects only (3) or metadata and digital objects (1).

23 aggregators (more than half) say that the licence they apply is transferable, and 29 allow them to give the metadata to Europeana but not for all uses: only 16 aggregators agree that the metadata could be reuse, for instance, for the virtual exhibitions. The reuse by third parties is much more restricted: only 12 aggregators agree upon.

The negative feedbacks could be due to the fact that the Europeana agreements are changing at the moment of this paper.

2.4 Identifiers

23 aggregators use one or more kind of persistent/permanent identifiers in their aggregation process (FIG. 10 and FIG. 11).

Many different international standards for persistent or permanent identifiers are applied. However, it is worth to stress the fact that 6 aggregators use own internal systems for the identification of the resources.

PERMANENT IDENTIFIERS	
ISBN	5
ISSN	5
DOI	3
HANDLE	2 + 2 under testing
UID	1
LSID	2 under testing
Internal codes (ID, inventory number, legal deposit code)	4

FIG. 10 – Use of permanent identifiers

PERSISTENT IDENTIFIERS	
PURL	3
URN	11
ARK	1
URI RDF	1
Permalinks	1
Own system based on OAI id specification	1
Internal codes	2
Other	3

FIG. 11 – Use of persistent identifiers

Persistent/permanent identifiers are quite commonly used by the providers of the aggregators: 13 aggregators even declared that over 50% of their content partners do it.

This result is definitely better as to the survey 2009 when a very low use of identifiers emerged (only 39% of the aggregators declared to apply them); now there is a growing awareness about the importance of identifying in a persistent way the digital resources.

2.5 Multilingualism

Translation of metadata

The aggregators may have metadata in multiple languages but this doesn't mean that they necessarily translate them; in fact, since many aggregators are European or international projects, they manage multilingual content by nature (10), even in 30 or more different languages like EuropeanaLocal (30 languages), BHL (32 l.), and The European Library (35 l.). Within EFG metadata can be translated into 13 languages because they are sometimes structured using multilingual controlled vocabularies.

As to national, regional, local aggregation initiatives, they mainly manage content in 1 or 2 languages (13 and 10 occurrences each); this means that metadata are usually in the native language and are seldom translated into English to reach a wider audience.

Multilingualism and portal public interface

19 aggregators have the portal interface only in one language (native language or German and English as working languages); however, 24 aggregators are used to providing the web pages of their portals into 2 or 3 languages at least. Best practices in the field are ECLAP (21 languages) and The European Library (35) (FIG. 12).

# OF LANGUAGES	# OF AGGREGATORS
1	19
2	8
3	8
4	1
5	1
>5	6

FIG. 12 – Amount of languages used in the website pages

English is obviously the language most frequently used for websites and translations (FIG. 13).

WEBSITE LANGUAGE	RECURRENCES
English	29
French	14
German	11
Dutch	10
Spanish	8

Italian	7
Czech	6
Danish	5
Finnish	5
Hungarian	5
Polish	5
Portuguese	5
Swedish	5
Estonian	4
Greek	3
Slovenian	3
Norwegian	2
Romanian	1
Russian	2
Serbian	2

FIG. 13 – List of the European languages used in the aggregators' web interface

Use of multilingual resources

The use of multilingual vocabularies for the aggregation process is not so widespread: only 19 aggregators are used to doing it (the same trend as Survey #1 of 2009).

Few aggregators hold multilingual resources like thesauri and online dictionaries (15, equally distributed between EU/international projects and national/regional/local initiatives) and even fewer are available as open source (2 upon request, 4 on the web sites).

Some example:

- **CulturaItalia**
http://www.culturaitalia.it/pico/thesaurus/4.2/thesaurus_4.2.0.skos.xml
- **The National Library of Finland**
<http://www.nationallibrary.fi/libraries/thesauri.html>;
<http://www.yso.fi/onki2/overview?o=http%3A%2F%2Fwww.yso.fi%2Fonto%2Fyso&l=en>
- **DISMARC**

<http://dev.ait.co.at/demo/thesauri/>

2.6 End-user services

Services provided by the aggregators

Similarly to Survey #1, portals in general offer the basic services like search and retrieve information, downloads, and the possibility of saving and sharing it; nevertheless, there is a significant improvement of the editorial services (newsletter and articles), feed RSS, and the possibility for the users to contribute the portal with UGC.

Advanced services also are more common in comparison with 2009, with particular reference to semantic search using linked data, print and digitisation on demand, and e-commerce, while the implementation of social networks is negligible.

PROVIDED SERVICES	# AGGR.
Portal (search, retrieve information)	35
Downloads	17
Saving, sharing searches and items	15
Newsletters	14
Editorial content	14
User generated content	12
Feed RSS	11
Semantic search (Semantic searching uses linked data, which gives machines the ability to make associations and put search terms into context)	6
Print on demand	6
Alert	4
Digitalisation on demand	4
Information, Training and technical support for the content providers	4
E-commerce	3

Export in XML	1
Social networking	1
Sending of electronic card via email	1

FIG. 14 – List of services provided by the aggregators

Monitoring the users

25 aggregators monitor the access to their portal; the remaining 18 don't do it or can't apply this question since they don't go on line publicly.

23 answers were gathered on the amount of unique visitors per month; the results are uneven: from 500-700 of Erfgoedplus to over 60,000 of Collections.

The average length of the visitor sessions is unequal too: 4 aggregators have basic results (less than 2 minutes), other 3 acceptable results (between 2 and 2 minutes 59 seconds), and even 10 have very good performances over 3 minutes (Culturegrid reaches 8 minutes, PIONEER 6 minutes and 14 seconds, The European Library 5 minutes and 15 seconds, Collections, DISMARC, and Manuscriptorium 5 minutes each, Archiefbank 4 minutes).

Google Analytics is the most used software for the elaboration of statistics about the aggregator web interface (14); other tools are AWStats (4), the log files analysis (2), Webalizer (1) as well as in-house software (2).

Many aggregators (26) don't require the users to be registered to access the content; others (13) ask for registration if the users want to enjoy advanced services (browsing specific content, upload files, create personal galleries, etc.).

Only 8 aggregators can give an overview of their publics: apart from Kulturarv Västernorrland and The European Library,

whose end-users are mainly experts and researchers, the other aggregators are used by all publics, mainly by the general public.

Periodical monitoring of the users is carried out by 7 aggregators only.

The aggregator users come from a large variety of geographical areas. Every aggregator has its own public: the user of local interest aggregators like Archiefbank, Erfgoedplus or Thuis in Brabant come from an homogeneous geographical area (Belgium and The Netherlands); by the other hand HISPANA reaches various Spanish mother tongue countries (Spain of course, but also Mexico and Argentina) (FIG. 15).

USERS PLACE OF ORIGIN	# AGGR.
Germany	14
USA	12
France	10
UK	10
the Netherlands	9
Belgium	6
Italy	6
Spain	3
Canada	3
Serbia	2
Austria	2
Finland	2
Poland	2
Japan	1
Slovenia	1
Croatia	1
Switzerland	1
Hungary	1
Malta	1
Mexico	1
Argentina	1
Czech Republic	1
Estonia	1
Russia	1
Sweden	1
Ukraine	1
Romania	1

FIG. 15 – Geographical provenance of the users

Promotion of the aggregator

The active promotion of the aggregator seems to be not one of the main activities of the aggregators since only 32 valid feedbacks on the specific question were received.

Similarly to the survey 2009, the main ways to promote the aggregators are the conference speeches (93.8% of the aggregators do that) and printing brochures and leaflets (75%). Alternative ways like search engines optimisation (53.1%) and the use of social networks (50%) significantly grew in comparison with the previous survey. Specific promotional events and links with other programmes are less used (64.9 and 34.4% each) to reach wider audiences.

By the other hand the aggregators' staffs seldom demonstrated to have a lively imagination and organised prize competitions to promote the portals; links to Wikipedia multilingual pages connected to specific topics or content of the portal is also a strategy as well as the aggregator introduction page on the free encyclopaedia.

Europeana end-user activities

The Europeana end users activities raise the interest of the aggregators' managers and 30 declared their availability to take part; the virtual exhibitions are preferred by almost all (90%), but also the user communities (76.7%) and the geolocation referring (63.3%) are chosen. However some aggregator highlighted also the existing obstacles to these activities like the financial contribution or the lack of ownership on the rights on digital resources.

2.7 Finance and sustainability

Funding

Usually the aggregators (20) combine budgets from various sources, public and private.

Funding from Ministries of culture or education totally supports the huge national aggregators or digital libraries and the portals of regional interests (9); however the national funding can be sometimes integrated by other budgets coming from one or two sources more, like European programmes, specific projects (CulturalItalia), Membership subscriptions (like for Manuscriptorium and Vlaamse Kunstcollectie), other bodies giving grants like foundations.

By the other hand, the sole European programmes cover mostly the aggregators feeding Europeana like BHL, DISMARC, EFG, EuropeanaLocal, and HOPE but also national initiatives like Kypriana and Letonica (the cross-domain national aggregators of Cyprus and Latvia), eSbirky (the aggregator of the Czech museum, born under the aegis of ATHENA).

Finally it is worth to mention that 13 aggregators benefit of financial resources that are allocated in the yearly budget of the institution; this information is particularly relevant since – either the funding is integrated by other sources or not – it demonstrates that the content aggregation process is part of the mission of the supporting institution.

The length of term for the funding is mostly 3-5 years (14) but also very short funding period are quite common (12). Archiefbank, Kulturpool and the National Digital Library of Finland are the best practices in the field since the length of their funding period is higher than 5 years (FIG. 16).

LENGTH OF FUNDING PERIOD	# AGGREGATORS
--------------------------	---------------

1 year	12
2 years	5
3-5 years	14
>5 years	3

FIG. 16 – Length of funding period

Budget breakdown

Only few aggregators managed to give a breakdown of their own budget. Staff expenses are the largest because they take up from 34 to 98% of the total budget, and sometimes IT equipment can reach 47% too. Marketing and promotion, overhead and facilities, and software licences don't weighing significantly on the budget.

18 aggregators have a business plan but only one published online¹!

The surveyed aggregators are initiatives that don't have the mission to generate the revenues: all respondents left out the specific question on this topic.

Staff

27 answers were received about the Full Time Equivalent (FTE) workers that aggregators employ for management and implementation.

Experts for the technical development and the ingestion procedure as well as the editorial staff are the people mainly involved into the aggregators building process.

Staff for marketing is reduced because – as seen before – aggregators invest little money on this activity (FIG. 17).

AREA	FTE workers
Technical development	55

¹ The National Digital Library of Finland at <http://www.kdk2011.fi/fi/raportit-ja-julkaisut/104-asiakasliittymaen-hyoetyanalyyysi> (in Finnish).

Ingestion staff	27
Editorial staff	26
Strategic and business development	24
Research and consulting	13
Marketing	10
Other	10
TOTAL FTE (17 aggregators)	165

FIG. 17 – Aggregators’ FTE workers

CulturaItalia and BHL have the wider staffs (19 and 16 FTE each).

In fact aggregators promote Europeana and related activities with speeches, printed material (flyers and brochures), mailing lists, expert meetings, on the website pages, every-day work with users, and articles in professional reviews; the Europeana Facebook Fan Page is also used. However Europeana is part of the communication plans of the aggregators funded by The European Commission because on of their goals is the content provision.

2.8 Other

Relation with Europeana

Owing to the plans of each aggregator, Europeana should receive over 37 M digital items between 2010 and 2012 (see table in ANNEX V)!

The Europeana Search API to search, retrieve and display Europeana results on the aggregator site is welcomed by 32 aggregators upon 43. The negative answers mainly concern the fact that the issue wasn’t taken into consideration yet.

Aggregators are pleased to be part of the Europeana building process: networking, sharing knowledge and services, extension of offered knowledge, staying up-to-date with the technical developments taking place in the digitisation domain (including discussions on multilingualism, metadata enrichment, tools and apps development), and, of course, wider visibility for the aggregator and its providers, are the most important reasons for being part of the Europeana family. Furthermore, Europeana is considered as a stimulus for digitization and documentation of local collections.

By the other hand, aggregators offer Europeana their expertise, prestigious masterpieces, visibility and promotion at national and local levels.

3. Conclusions

The aggregators that answered the ATHENA and Europeana questionnaire share some features:

- they are mainly funded with public money coming from governmental bodies (ministries of culture or/and education, above all) and European programmes;
- although only few of them (less than 1/3) can count on a allocated yearly budget for the aggregation, the funding time span is becoming longer in comparison to Survey #1. This means that they are getting even more structured and embedded into the policy of the promoting institutions;
- they make digital resources circulate. It is common that some aggregators deliver Europeana content of other aggregators (e.g. The European Library delivered content of Digital Library of Serbia, Digital Library of Slovenia, Letonica, and Manuscriptorium);
- they are able to manage at once many different kind of metadata, although Dublin Core is still the most diffused schema;
- they often act as repositories of digital content for their content providers;
- they provide basic services of search and retrieve information, downloads, newsletter, and editorial content. However semantic search and publication of UGC are increasing;
- they improved the use of persistent or permanent identifiers and their content providers too;
- they pay moderate attention to multilingualism. The translation of the web interface is often careful, but the use of multilingual thesauri and controlled vocabularies is not so common;
- licences or specific agreements that regulate use and reuse of (meta)data are not always applied; this could generate some critical issue for Europeana;
- they participate in Europeana and wish to largely contribute.

ATHENA can interact and contribute to the landscape of aggregators with the outcomes of the project in order to assure, beyond the sustainability of the aggregated data in a short term perspective¹, also the sustainability of the standards, publications, and ideas that were elaborated during 30 months.

LIDO is one of the major ATHENA outcomes since it is a valid support for standard harmonisation; it is now well known and applied by some aggregators (p. 15), as well as by some ATHENA content providers within their national projects.

ATHENA produced technical publications (based on the deliverables) on different topics; many European aggregators could benefit from them, in particular the recommendations about persistent identifiers².

Licences application, in particular for regulating the reuse of content by Europeana and third parties, could be supported with the use of the ATHENA IPR step-by-step guide, the most straightforward way of clearing rights in order to achieve the legal basis for the resource exploitation³.

¹ This will happen within the Linked Heritage project (see deliverable 1.5 ‘Sustainability concept’ p.7).

² <http://www.athenaeurope.org/getFile.php?id=779>

³ http://devel.silktech.gr/athenaeurope_ipr/lang_en/page/home-page

Furthermore, ATHENA actively participated to the definition of the Europeana licensing framework, with particular regards to the recent Data Exchange Agreement that allow the commercial used of metadata in order to prevent some museums' reluctance¹.

ATHENA proposed the ATHENA thesaurus to foster multilingualism in content aggregation. The ATHENA Thesaurus is a network of controlled vocabularies, that is, an amount of terms organised by domains of description and structured thanks to bridges in-between. This ATHENA Thesaurus is: SKOSified (and fits with Europeana requirements), free of right, evolving, available online².

Europeana is now investigating how to deal with the end of aggregation projects like ATHENA, MIMO, Judaica etc.; a suggestion for further investigation about aggregators could concern the way they may maintain their repository or their data in Europeana, like ATHENA did with the help of another project, Linked Heritage, involving a consistent part of the ATHENA consortium

It is clear that aggregators are now became more than projects, but a strategic model for the long term sustainability of digitisation across Europe.

¹ See D1.5 'Final report', WP6 activities, for all details concerning the Data agreement workshop for ATHENA partners organised in Bruxelles in cooperation with Europeana on April 8, 2011.

² http://www.athenaeurope.org/athenawiki/index.php/Athena_Thesaurus

Annex I: The questionnaire



WELCOME

Welcome to the Survey of Aggregators. You are asked to participate as a known aggregator of cultural heritage content.

Europeana.eu brings together digitised content from Europe's cultural and scientific heritage organisations, and makes that content accessible to Europe's citizens and to the wider world. Europeana works mainly with aggregators of content – cross and single domain. National cross domain aggregators include Hispana, who brings together content from a range of Spanish heritage organisations; single domain aggregators include The European Library, who aggregate content from Europe's national libraries.

Europeana and ATHENA projects launch the Survey for Aggregators with the purpose to verify the shared issues and enable the establishment of aggregators in contributing content to Europeana. The Survey wants also to investigate possible relations, similarities and parallelisms with other projects aimed at content provision.

ABOUT THE SURVEY

1. Purpose of the survey

This is the second Survey of Aggregators. Its purpose is to identify shared issues and help in the establishment of aggregators to contribute content to Europeana.eu directly or through satellite projects.

As a follow up of the first Survey it will update information on the state of aggregation across Europe. The results will suggest the strategies, activities and services that need to be developed to help aggregators act as portals in their own right and help in the production of Europeana.

2. Who needs to supply the answers to the survey?

The person in charge of Strategy or IT/digitisation in the organization. S/he can enlist the help of others, such as metadata and technical advisors.

3. How is the survey structured?

This survey deals with the main areas of establishing a partnership between aggregators and Europeana. The survey consists of questions following these main subjects:

1. General information
2. Content Ingestion
3. Partner Relations
4. Identifiers
5. Multilingualism
6. End-User Services
7. Finance and sustainability
8. Others

4. Deadline & contacts

We kindly ask you to provide us with your responses before 1st of November 2010. If you have any questions to the survey please don't hesitate to contact us.

Email: feedback@europeana.eu, cc annette.friberg@kb.nl;
marzia.piccininno@beniculturali.it; v.vassallo@cyi.ac.cy

www.Europeana.eu

Thank you very much for your time and help.

.....

GENERAL INFORMATION

1. Name of aggregator and its location
2. Primary contact
3. Technical contact (if possible)
4. Please specify the geographical level of aggregation. Geographic level is the level at which your organisation will operate in terms of the data it collects i.e. whether the sources are regional, European or worldwide. An example of an organisation operating at the European geographic level is Europeana, which collects data located around Europe.
 - a. Worldwide (EU countries and beyond)
 - b. European (only EU countries)
 - c. National
 - d. Regional
5. Please specify type of Aggregation
 - a. Cross domain - working across several sectors and collecting material from museums, libraries and archives. (Examples: CultureGrid, CulturaItalia, Collections.fr and Europeana)
 - b. Single (=Vertical) domain representing and collecting data from one sector only such as museums, libraries or archives (Example: The European Library)
 - c. Thematic Aggregation – bringing together content on a specific theme from across several sectors or one sector. (e.g. Judaica Europeana, collecting material from multiple sectors about Jewish culture)
6. Is your content already online in your own portal?
 - a. Yes
 - b. No
7. Have you already delivered aggregated content to Europeana?
 - a. Yes
 - b. No
8. If Yes:
 - a. directly to Europeana
 - b. through other European projects or aggregator
9. How many digital objects have you aggregated to date? Europeana counts the number of records. So a document with 300 pages through which you can browse on the local site but which has metadata only at the level of the document will count as one digital object and one record in Europeana. Pages from a book that have separate meaningful metadata useful for finding the individual page may be counted as separate digital objects and records. Museums collecting physical objects usually consider the description of a physical object, together with one or more images attached, to be the

digitized version of that object. Such descriptions are therefore considered as digital objects and records in the context of Europeana.(A set of digital objects that is not a digital object itself, but has meaningful metadata of its own, can also be considered as a record). Total number of objects:

10. What is the percentage of the total for each type of digital object?
 - a. Audio
 - b. Video
 - c. Text
 - d. Image
 - e. Other
11. How many digital objects do you expect to aggregate within the next year and from which sources? NUMBER - SOURCE (Example: 12450 - Archive of...) Please, one answer per row.
12. Year of launch of aggregator
13. Do you store and maintain full digital content (e.g. full size images, pdf etc)?
 - a. Yes
 - b. No
 - c. Other
14. If no, do the metadata provide direct links to the digital objects?
 - a. Yes
 - b. No
15. Please list aggregators you are aware of in Europe Specify Nation, if Cross Domain, Vertical or thematic, Name and URL (Example: Italy, Cross Domain, CulturalItalia, <http://www.culturaitalia.it>) Please, list one per row.
16. Is your aggregator supported by a larger organisation?
 - a. Ministry
 - b. Foundation/charity
 - c. Museum
 - d. Library
 - e. Archive
 - f. Other
17. How would you position yourself in relation to Europeana and other Aggregators and Portals?
18. Do you have an estimate of the total digital cultural heritage objects in numbers to be aggregated broken down by: libraries, museums, archives and audiovisual? (Example: Libraries: 12300; Museum: 4520; Archives: 34000; Audiovisual: 50). Please, specify the numbers and until when.

CONTENT INGESTION

19. Do you have a content strategy or collections development policy?
 - a. Yes
 - b. No
 - c. In development
20. Can it be made available?
 - a. Yes
 - b. No
21. To what area of content do you provide access?
 - a. Art
 - b. Humanities
 - c. Social sciences

- d. Scientific
 - e. Other
22. To what type of content do you provide access?
- a. Books
 - b. Journals
 - c. Newspapers
 - d. Datasets
 - e. E-These
 - f. Catalogues
 - g. Pictures
 - h. Photos
 - i. Video
 - j. Audio
 - k. 3D
 - l. Other
23. What metadata schema do you use to aggregate your collections?
- a. Dublin Core
 - b. EAD
 - c. Museum.dat
 - d. CIDOC CRM
 - e. LIDO
 - f. INSPIRE
 - g. GML
 - h. X3D
 - i. Other
24. How do you aggregate/harvest (meta)data?
- a. OAI-PMH
 - b. FTP
 - c. Opensearch
 - d. SRU
 - e. Other
25. Are you able to export for reuse, content you have aggregated?
- a. Yes
 - b. No
26. If yes, please indicate your export mechanisms.
- a. OAI-PMH
 - b. FTP
 - c. Opensearch
 - d. SRU
 - e. API
 - f. Other
27. Are you able to export for reuse, metadata you have aggregated?
- a. Yes
 - b. No
28. If yes, please indicate your export mechanisms.
- a. OAI-PMH
 - b. FTP
 - c. Opensearch
 - d. SRU
 - e. API

- f. RSS FEED
 - g. Other
29. Is your system based on free software, open source software, proprietary software (for definition see <http://www.gnu.org/philosophy/categories.html>).
- a. Free software
 - b. Open source software
 - c. Proprietary software
30. Are you aware of software tools for the ingestion of data (e.g. ATHENA ingester or Europeana content checker)?
- a. Yes
 - b. No
31. Do you use one of these software tools to ingest data?
- a. Yes
 - b. No
32. Do you have an ingestion plan with scheduling and amounts?
- a. Yes
 - b. No

PARTNER RELATIONS

33. For how many institutions do you currently aggregate content?
34. Do you actively approach potential content providers?
- a. Yes
 - b. No
35. If Yes, what are your selection criteria for creating new partnerships with content providers?
- a. Importance of the institution
 - b. Development of technical functionality
 - c. Service and continual supply of content
 - d. Relevance of the collection(s) to your collection development policy
 - e. Topic of the collection(s)
 - f. Other
36. Do you have a licence or agreement for the publication of the data on your site?
- a. Yes
 - b. No
37. If yes, please provide a copy or, if it is a published model licence (e.g. Creative Commons, Europeana).
- a. CC0 (Creative Commons)
 - b. CC by SA (Creative Commons by Share Alike)
 - c. CC by SA non Commercial
 - d. Other
38. If not, how do you regulate the relationship with the content providers?
39. The licence is applied to
- a. Metadata
 - b. Digital content
 - c. Thumbnails
40. Is the licence transferable?
- a. Yes
 - b. No
41. Will the licence allow you to give the metadata to Europeana?
- a. Yes

- b. No
- 42. Will the licence allow Europeana to re-use the content for own purposes for example featuring your content in a Virtual Exhibition?
 - a. Yes
 - b. No
- 43. Will the licence allow re-use of the content by third parties?
 - a. Yes
 - b. No

IDENTIFIERS

- 44. Do you as an aggregator use persistent or permanent identifiers?
 - a. Yes
 - b. No
- 45. If permanent, which one?
 - a. ISBN
 - b. ISSN
 - c. DOI
 - d. HANDLE
 - e. Other
- 46. If persistent, which one?
 - a. PURL
 - b. URN
 - c. ARK
 - d. Other
- 47. What percentage of your content providers uses persistent identifiers?
 - a. 0-10%
 - b. 10-30%
 - c. 30-50%
 - d. Over 50%

MULTILINGUALISM

- 48. Do you use multilingual vocabularies?
 - a. Yes
 - b. No
- 49. How many languages is your metadata in?
- 50. 50. If you have collection descriptions, please indicate the number of languages and the top 10 they are in
- 51. Is the metadata translated?
 - a. Yes
 - b. No
- 52. In which languages is your portal interface available?
 - a. Only native language
 - b. English
 - c. German
 - d. Others
- 53. Is the interface totally translated into other language(s) or only part of it (please indicate percentage)?
- 54. Do you hold multilingual resources (e.g.: thesauri, online dictionaries)?
 - a. Yes

- b. No
55. Are these available as open source?
- a. Yes
 - b. No

END-USER SERVICES

56. If you have a portal what kind of services does it offer?
- a. E-commerce (buy a book, etc.)
 - b. Portal (search, retrieve information)
 - c. Newsletters
 - d. Alert
 - e. RSS FEED
 - f. Semantic search (Semantic searching uses linked data, which gives machines the ability to make associations and put search terms into context)
 - g. Print on demand
 - h. Digitalisation on demand
 - i. Downloads
 - j. Saving, sharing searched and items
 - k. Editorial content
 - l. User generated content
 - m. Other
57. How many unique visitors (individual visitors who may make multiple visits) does the site attract per month?
58. What is the average time spent on the site?
59. Do you monitor the access to your portal?
- a. Yes
 - b. No
60. How do you promote your portal?
- a. Search Engine optimisation
 - b. Linking programme
 - c. Speeches at conferences
 - d. Arrange events specifically to promote the portal
 - e. Traditional promotional material (brochure, leaflets, bookmarks, etc.)
 - f. Dissemination through social networks (Facebook, Twitter, etc.)
 - g. Other
61. Do your users need to register?
- a. Yes
 - b. No
 - c. Optional
62. Please provide an estimate of your end-users in percentages of your total end-user database?
- a. Students
 - b. Researchers
 - c. Professional
 - d. General public
 - e. Experts
 - f. Other (please specify)
63. Do you periodically monitor the user profile?
- a. Yes
 - b. No

64. What are the 5 top geographical origins of these users?
65. Would you like to participate in Europeana's end-user activities?
- Road shows
 - Communities
 - Virtual exhibitions
 - Geolocation referring

FINANCE AND SUSTAINABILITY

66. Where do you get your funding (you may indicate more than one)?
- Membership subscriptions
 - Ministries of culture or education
 - Grant giving bodies (e.g. charities, foundations)
 - European funding. Please specify
 - Other funding (national, regional, private etc.) Please specify
 - Allocated in the yearly budget of the institution(s)
67. What is the length of term for the funding?
- 1 year
 - 2 years
 - 3-5 years
 - > 5 years
68. Could you give a breakdown of your budget?
- Staff
 - Marketing, promotion
 - Overhead, facilities
 - IT equipment
 - Software
 - Other
69. Do you have a business plan or other management reporting?
- Yes
 - No
70. If your portal or aggregation activity generated revenue in 2009, please indicate the breakdown of your own (self generated) revenues in 2009 and if possible the value received in subsidy. Please ensure percentages add up to 100%.
- Self generated revenue
 - From private/individual donations
 - from commercial income (e.g. fees, subscriptions product sales, advertising)
 - from sponsors
 - from other sources (please specify)
71. How many Full Time Equivalent workers relates to the management and implementation of the aggregator?
- Strategic and Business Development
 - Technical development
 - Editorial staff
 - Research and consultancy
 - Ingestion staff
 - Marketing and promotion
 - Other

OTHER

72. How many digital objects approximately do you plan to deliver to Europeana?
- a. 2010
 - b. 2011
 - c. 2012
73. Would you be interested in using the Europeana Search API to search, retrieve and display Europeana results on your own site?
- a. Yes
 - b. No
74. If you would not wish contribute to Europeana please say why?
75. What are your expected benefits from contributing and being part of Europeana?
76. And what benefits will Europeana gain with your contribution?
77. How do you promote Europeana?

Annex II: Aggregators' list

(in **yellow** the aggregators that participated in the Survey for Aggregators #1 in 2009)

Name	Location	Level of Aggregation	Domain level	URL
Archiefbank	Belgium	Worldwide	Single domain (archive)	www.archiefbank.be
BAM	Germany	National	Cross-domain	www.bam-portal.de
BHL-Europe	Germany	Worldwide	Cross-domain; Thematic (biodiversity)	Not yet (august 2011). http://www.bhl-europe.eu/ (project website)
Carmentis	Belgium	Regional	Single domain (museum)	http://carmentis.kmkg-mrah.be/eMuseumPlus http://master.kmkg-mrah.be/
CER.ES	Spain	National	Single domain (museum)	http://ceres.mcu.es/
Collections	France	National	Cross-domain	http://recherche.culture.fr
CulturaItalia	Italia	National	Cross-domain	www.culturaitalia.it
Culture Grid	United Kingdom	National	Cross-domain	www.culturegrid.org.uk
Digital Libraries of Serbia	Serbia	National	Single domain (library)	http://digitalna.biblioteka.rs/
Digital Library of Slovenia	Slovenia	National	Cross-domain	www.dlib.si
DISMARC	Germany	Worldwide	Thematic (music and music related items)	www.dismarc.eu
ECLAP	Italy	Worldwide	Thematic (performing arts)	http://www.eclap.eu
EFG (The European Film Gateway)	Germany	European + Switzwerland	Vertical domain (European film archives, film museums and cinémathèques)	Not yet (Spring 2011). http://www.europeanfilmgateway.eu/ (project website)
Encycloasturias - Mediateca	Spain	Regional	Cross-domain	http://mediateca.uniovi.es

de la Universidad de Oviedo				
Erfgoedplus.be	Belgium	Regional	Cross-domain	www.erfgoedplus.be
eSbirky	Czech republic	National	Single domain (museum)	www.esbirky.cz
EuropeanaLocal Austria	Austria	National	Cross-domain	http://www.europeana-local.at
Expo @ AcrossLimits	Malta	National	Thematic (private collections)	http://expo.acrosslimits.com
Gateway Bayern	Germany	Worldwide	Single domain (library)	http://www.digitale-sammlungen.de/index.html?c=digitale_sammlungen&l=de
Hellenic Aggregator for Europeana	Greece	National	Cross domain	http://aggregator.libver.gr (not available as public portal)
Hispana	Spain	National	Cross domain	http://hispana.mcu.es
HOPE project	The Netherlands	Worldwide	Thematic (social history, labour history)	http://www.peoplesheritage.eu/ (project web site)
Judaica Europeana	France/UK	Worldwide	Thematic (Jewish participation in the development of urban culture in Europe)	http://www.judaica-europeana.eu (project's web site)
Kulturarv Västernorrland	Sweden	Regional	Cross-domain	http://kulturarvvasternorrland.se/
Kulturpool	Austria	National	Cross-domain	www.kulturpool.at
Kypriana	Cyprus	National	Cross-domain	www.kypriana.eu
Letonica	Latvia	National	Cross-domain	www.lndb.lv
Manuscriptorium	Czech Republic	Worldwide	Single domain (written cultural heritage)	http://www.manuscriptorium.com
MIMO (Musical Instruments)	United Kingdom	European	Thematic domain (Musical)	Not yet. http://www.mimo-project.eu/

Museum Online)			Instruments)	(project website)
MovE	Belgium	Regional	Single domain (museum collections)	www.museuminzicht.be
Muis	Estonia	National	Cross domain	www.muis.ee
National Digital Library of Finland	Finland	National	Cross-domain	http://www.kdk2011.fi/en (project web site, not yet available for the public)
PIONIER Digital Libraries Federation	Poland	National	Cross-domain	http://fbc.pionier.net.pl/
Porstua Web Service	Finland	Regional	Single domain (library)	http://www.porstuakirjastot.fi/
Portail de la musique contemporaine	France	National	Thematic (contemporary music)	http://www.musiquecontemporaine.fr
Public Library - Varna	Bulgaria	Regional	Cross-domain	http://catalog.libvar.bg/search-DC.html
Registo Nacional de Objectos Digitais - RNOD	Portugal	National	Single domain (library)	Not yet
SCRAN	UK (Scotland)	National	Cross-domain	www.scran.ac.uk
SOCH	Sweden	National	Cross-domain	http://www.ksamsok.se/in-english/
The European Library	The Netherlands	Worldwide	Single domain (library)	www.theeuropeanlibrary.org
Thuis in Brabant	The Netherlands	Regional	Cross-domain	www.thuisinbrabant.nl
Vlaamse Kunstcollectie	Belgium	Regional	Single domain (museum)	www.vlaamsekunstcollectie.be
www.africamuseum.be		Worldwide		

Annex III: Positioning

Answers to question 17: “How would you position yourself in relation to Europeana and other Aggregators and Portals?”

AGGREGATOR	RESPONSE
Archiefbank	As a potential provider through APENet
BAM Portal	Major portal for Germany
BHL - Europe	Although we are probably not one of the largest aggregators in terms of numbers, we deliver science content to Europeana that is underrepresented otherwise. With our global network of partners we get experience and knowledge that makes us special. We are well integrated in the established bioinformatics infrastructure with numerous projects and initiatives all over the world
Carmetis	Direct partner of Europeana
CER.ES	Ceres provides a large number of records (more than one hundred thousand) to Hispana. Ceres is working hard to provide new digital contents to Hispana and, afterward, to Europeana. We are very interested in participating in networks of dissemination of cultural heritage. We wish to spread the rich heritage that is preserved in Spanish museums
Collections	Collections.fr is one of the three national aggregators contributing to Europeana. Collections.fr is the national cross-domain aggregator for France, the thematic aggregators from INA (Institut National de l'Audiovisuel) and BnF (Bibliothèque nationale de France) are also contributing to Europeana
CulturaItalia	CulturaItalia is the Italian national aggregator
Culture Grid	UK national aggregator to Europeana Potential provider and consumer of services for other portals
Digital Libraries of Serbia	Linked open data to all aggregators, and Europeana too
Digital Library of Slovenia	As national cross domain portal
EFG	Aggregator for film and film-related material (audiovisual sector)
Encycloasturias - Mediateca de la Universidad de Oviedo	We started this year; we would encourage the participation of universities in the project. The Universiade can add cultural content, scientific and educational, its closest regional environment
Erfgoedplus.be	Erfgoedplus.be is a fully operational aggregator and portal for really local content with a strong bottom-up approach, in support of heritage collections of two provinces. It can expand to other provinces or institutions in or outside of Belgium. To the local institutions or communities it offers also services as repository for their content. Europeana is an additional channel of access for the content.
eSbirky	National museum is partner of ATHENA project. At the present time we are working on a new on-line application www.esbirky.cz (eCollections), which was intended from the beginning to simplify the contribution of Czech museums to Europeana. We are looking forward that this website will attract many others Czech memory institutions.

EuropeanaLocal	We are an aggregator for regional and local content all over Austria
Expo @ Acrosslimits	National Aggregators for Malta
Gateway Bayern	Bavarian State Library is the aggregator for libraries within the Bavarian Regional Network.
Hispana	As a national aggregator
HOPE project	Independent Aggregator that both feeds into Europeana and other channels
Kulturav Västernorrland	Regional distributor to national aggregators (Ksamsök, Libris, NAD) and to international aggregators as Europeana
Kulturpool	National Cross-domain Aggregator and Search-Portal with group specific services
Kypriana	Content Provider and Aggregator
Letonica	Currently mainly as a partner in TEL and thematic aggregators, hopefully in future - a national aggregator or at least a national library sector aggregator to Europeana - if there's a local political support
MovE	Small aggregator focussed on local content
Muis	Small, but enlarging aggregator
National Digital Library of Finland	We will be both the national portal for end users and the national aggregator of Europeana Finnish libraries, archives and museums
PIONIER Network Digital Libraries Federation	We position the Polish Digital Libraries Federation as the best and most complex source of metadata of digital objects available on-line from Polish scientific and cultural institutions. We cooperate with Europeana and several other aggregators in this subject (eg. DART-Europe and ViFaOst). We also position the Federation as the source of new tools, services and possibilities for our metadata providers
Porstua Web Service	Aggregator of regional information of Finnish Ostrobothnia. Among other provincial libraries.
Portail de la musique contemporaine	Rare material: free access to significant audio excerpts of contemporary - i.e., recent - art music under copyright
RNOD	RNOD aims to aggregate Portuguese digital content to Europeana
Scran	National aggregator for Scottish culture and heritage
SOCH	National aggregator for immobile heritage and museums
The European Library	Alongside the other single-domain portals
Vlaamse Kunstcollectie	data deliverer, nothing more, nothing less
www.africamuseum.be	As an institutional data provider, via projects and playing a role to have African related content in EUROPEANA

Annex IV: Content providers of each aggregator

Archiefbank	100
BAM Portal	9
BHL - Europe	39
Carmetis	4
CER.ES	67
Collections	40
CulturaItalia, The Italian Culture Portal	30
Culture Grid	100
Digital Libraries of Serbia	25
Digital Library of Slovenia	30
DISMARC	65
ECLAP - European Collected Library of artistic performance	17
EFG - European Film Gateway	15
Encycloasturias - Mediateca de la Universidad de Oviedo	16
Erfgoedplus.be	160
eSbírky	6
EuropeanaLocal	17
Expo @ Acrosslimits	15
Gateway Bayern	1
Hellenic Aggregator for Europeana	8
Hispana	132
HOPE project	13
Judaica Europeana	12
Kulturarv Västernorrland	5
Kulturpool	9
Kypriana	4
Letonica	50
Manuscriptorium	100
MIMO-DB	10
MovE - Musea Oost-Vlaanderen in Evolutie	40
Muis	51
National Digital Library of Finland	0
PIONIER Network Digital Libraries Federation	300
Porstua Web Service	0

Portail de la musique contemporaine / Contemporary Music Portail	31
Public Library - Varna	6
Registo Nacional de Objectos Digitais - RNOD	0
Scran	300
SOCH	41
The European Library	48
Thuis in Brabant	12
Vlaamse Kunstcollectie	3
www.africamuseum.be	3
TOTAL	1,934

Annex V: Planned digital items to be delivered to Europeana

AGGREGATOR NAME	2010	2011	2012
Archiefbank	-	-	-
BAM Portal	8,000	-	-
BHL - Biodiversity Heritage Library for Europe	87,000	88,000	110,000
Carmetis	12,000	41,000	-
CER.ES	-	30,000	30,000
Collections	1 M	2 M	3 M
CulturaItalia, The Italian Culture Portal	-	-	-
Culture Grid	600,000	1,000,000	2,000,000
Digital Libraries of Serbia	10,000	12,000	15,000
Digital Library of Slovenia	76,000	200,000	200,000
DISMARC	100,000	100,000	
ECLAP - European Collected Library of artistic performance	0	400,000	1,000,000
EFG - European Film Gateway	300,000	300,000	
Encycloasturias - Mediateca de la Universidad de Oviedo	500	3,000	5,000
Erfgoedplus.be	50,000	25,000	25,000
eSbírky	20,000	500,000	100,000
EuropeanaLocal	30,000	180,000	280,000
Expo @ Acrosslimits	35,000	45,000	100,000
Gateway Bayern	80,000	80,000	80,000
Hellenic Aggregator for Europeana	-	-	-
Hispana	1,500,000	2,000,000	2,500,000
HOPE project	-	-	600,000
Judaica Europeana	0	300,000	200,000
Kulturarv Västernorrland	70,000	28,000	0
Kulturpool	0	80,000	150,000
Kypriana	1	3,000	3,000
Letonica	5,000	15,000	200,000
Manuscriptorium	40,000	40,000	100,000
MIMO-DB	-	-	-
MovE - Musea Oost-Vlaanderen in Evolutie	-	50,000	-
Muis	0	4,000	0
National Digital Library	-	-	-
PIONIER Network Digital Libraries Federation	450,000	600,000	750,000
Porstua Web Service	-	-	-
Portail de la musique contemporaine /	-	-	-

Contemporary Music Portail			
Public Library - Varna	-	-	-
Registo Nacional de Objectos Digitais - RNOD	-	60,000	-
Scran	360,000	-	-
SOCH	1,200,000	2,000,000	3,000,000
The European Library	2,000,000	500,000	4,500,000
Thuis in Brabant	-	-	-
Vlaamse Kunstcollectie	-	1,000	1,000
www.africamuseum.be	50,000	100,000	100,000
TOTAL	8,083,502	10,785,000	19,049,000