

DELIVERABLE

Project Acronym: EAGLE
Grant Agreement number: 325122
Project Title: Europeana network of Ancient Greek and Latin Epigraphy

Ingestion plan

Deliverable 3.2

version: 1.0

Revision: final

Authors:

Valentina Vassallo (The Cyprus Institute)

Contributors:

Silvia Evangelisti (UNIROMA1)
Anita Rossi (UNIBA)
Pietro Liuzzo (UHEI)
Christian Uhlir (PLUS)
Marjeta Šašel Kos (ZRC SAZU)
Anja Ragolič (ZRC SAZU)
Valentina Zović (UNIPU)

Reviewers:

Claudio Prandoni (Promoter)
Manuele Buono (Promoter)
Franco Zoppi (CNR-ISTI)

Project co-funded by the European Commission within the ICT Policy Support Programme		
Dissemination Level		
P	Public	X
C	Confidential, only for members of the consortium and the Commission Services	

Revision History

Revision	Date	Author	Organisation	Description
0.1	21/03/2014	Valentina Vassallo	The Cyprus Institute	First version
0.2	31/03/2014	Valentina Vassallo	The Cyprus Institute	Final version including the comments received from the reviewers
1.0	03/04/2014	Claudio Prandoni Antonella Fresa Pietro Masi Manuele Buono	Promoter	Formal check and improvements

Statement of originality:

This deliverable contains original unpublished work except where clearly indicated otherwise.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

TABLE OF CONTENTS

LIST OF ABBREVIATIONS.....	4
EXECUTIVE SUMMARY	5
1. INTRODUCTION.....	6
1.1 OBJECTIVES OF THE WP3 ‘METADATA MODEL, MAPPING AND INGESTION’	6
1.2 EAGLE AND THE EUROPEANA SISTERS PROJECTS.....	7
1.3 THE ROLE OF THE DELIVERABLE 3.2 IN THE PROJECT.....	8
2. CONTENT PROVIDERS AND DIGITAL COLLECTIONS	10
1.4 CONTENT PROVIDERS.....	10
1.4.1 <i>The enlargement of the Consortium: new Content Providers</i>	11
1.5 EAGLE COLLECTIONS.....	12
TABLE III. DESCRIPTION OF THE COLLECTIONS PROVIDED BY EAGLE CONTENT PROVIDERS.....	14
3. THE INGESTION PLAN	15
1.6 THE INGESTION PLAN FORM	15
1.6.1 <i>The first Survey (March 2013)</i>	16
1.6.2 <i>The second Survey (January 2014)</i>	18
1.7 DATA EXCHANGE AGREEMENT (DEA)	21
1.7.1 <i>Europeana Rights Statements</i>	22
1.8 CONTENT PROVIDERS PLANNING FOR THE DELIVERY OF EAGLE CONTENT TO EUROPEANA.....	23
1.9 WP3 PLANNING FOR THE CONTENT PROVISION.....	26
1.9.1 <i>Mid-term internal milestones</i>	27
1.10 MONITORING AND SUPPORTING THE PROVIDERS	28
1.10.1 <i>Direct contacts</i>	28
1.10.2 <i>Training sessions and demonstrations</i>	30
1.10.3 <i>Setting-up of a Support Forum</i>	31
1.10.4 <i>Relations with the other Work Packages</i>	33
1.11 FROM EAGLE TO EUROPEANA: THE INGESTION WORKFLOW.....	33
4. CONCLUSIONS	38
BIBLIOGRAPHY.....	39
ANNEX I	40
ANNEX II.....	45
1. <i>Provider information</i>	45
2. <i>Content description</i>	45
ANNEX III.....	47

LIST OF ABBREVIATIONS

ABBREVIATION	EXPLANATION
DoW	Description of Work
EDM	Europeana Data Model
WP	Work Package
WG	Working Group
DEA	Data Exchange Agreement

EXECUTIVE SUMMARY

This Deliverable, created by the Cyprus Institute (Cyl), with some contributions of the Content Providers opportunely cited in the text (see paragraph 3.3 of this Deliverable), reports about the planning and preparation of the ingestion in order to set up the workflow and the framework for the publication of the digital collections of the EAGLE Consortium in Europeana.

The document is structured in 4 Chapters plus 3 Annexes: an introduction about the scope of this Deliverable and its role in the project and in the Europeana environment; a section dedicated to the content and the collections to be published on the Europeana portal. A part dedicated to the implementation of the ingestion plan and of the activities to carry out the tasks of monitoring, supporting and gather feedback from the Content Providers. Finally, the part dedicated to the passages that describe the ingestion workflow from EAGLE to Europeana.

The Introduction of this Deliverable explains the aims of the Work Package and its role in the project. The first Chapter is an overview of the objectives of WP3: it describes the tasks of the WP3 and its overall objectives, focussing also on its Milestones. The purpose of the document and the relation of the work with other Europeana sisters' projects are explained.

Chapter 2 describes the EAGLE Consortium and the Content Providers that are involved in. Furthermore, it is explained the activity of enlarging the network through the involvement of new providers and affiliated partners. A particular attention is dedicated to the description of the EAGLE collections and their importance in the Europeana environment.

Chapter 3 is the core of the WP3 work because it illustrates the ingestion plan and the EAGLE data workflow to Europeana. One important task is the assessment of the amount of data to deliver to Europeana, respect to the content declared in the Description of Work. The creation of the ingestion plan form helps to report the available material and the amount of the content ingested during the project time span. The assessment, carried out through surveys that provide detailed information about the digital collections the partners will contribute, is reported in this section. Besides the results of the surveys, the report of the current amount of data and the related motivation, Chapter 3 discusses the importance of the signature of the Europeana Data Exchange Agreement for the publication of the data in Europeana and about the digital objects showed in the portal. The Chapter deepens reporting about the ingestion plan activities carried out by the Content Providers for the delivery of the content to Europeana through some examples, and the planning activities organized by WP3 for the overall content provision. Particularly for this part, the deadlines (mid-term internal milestones forms) and the methodologies (direct communications, monitoring plan, online support forum, training sessions and training material, WPs monitoring) set up to monitor and support the work of the Content providers are reported. Finally, the last part of the Chapter is dedicated to the description of all the steps of the workflow and guidelines to follow for the provision of content to Europeana.

Chapter 4 draws the conclusions and gives information about the next steps of the Work Package 3.

1. INTRODUCTION

1.1 OBJECTIVES OF THE WP3 'METADATA MODEL, MAPPING AND INGESTION'

WP3 'Metadata model, mapping and ingestion' is the core of the EAGLE project since it has to deal with the contribution of quality metadata and content from the individual institutional partners to the EAGLE Aggregation and Image Management (AIM) infrastructure implemented within WP4 ('EAGLE Aggregation and Image infrastructure'). In the WP3, the common EAGLE metadata structure to ensure full interoperability with Europeana has been defined (see Deliverable D3.1 'EAGLE metadata model specifications'). This will be the basis for the local mappings. The work within the Work Package addresses all the specific activities that are needed locally to harmonize the content coming from the different sources and archives of the project partners and prepare it for the ingestion process. The work consists of the implementation of the agreed standards and protocols, the preparation of the necessary mappings for the metadata conversions, the ingestion plan itself and finally the publication of the content into Europeana portal, including appropriate disambiguation. The role of the Content Providers is that one of checking, enrich and contextualize their metadata thanks to the data curation developed within WP4.

The Cyprus Institute, Cyprus, is the leader of the Work Package and leader of the tasks 'Planning and preparing the ingestion' and 'Implementing and monitoring the ingestion'. Other 3 Task leaders are associated: UAH, Spain ('Definition of the EAGLE metadata model' and 'Metadata mapping'); K.U.LEUVEN, Belgium ('Duplicates identification') and UNIROMA1, Italy ('Metadata enrichment and contextualization'). All the Content Providers contribute to the mapping and the ingestion of their content to the EAGLE Aggregator and to Europeana.

The overall objectives of the Work Package 3 are:

- To achieve the large scale implementation of agreed standards and best practices at the local and institutional level across the Best Practices Network (BPN)
- To define the common EAGLE metadata structure to enable BPN metadata harmonization
- To define mapping from local BPN metadata structures to the common EAGLE metadata structure
- To define mapping from EAGLE metadata structure to Europeana Data Model (EDM), to ensure full interoperability with Europeana.
- To ingest content into EAGLE and Europeana
- To check, enrich and contextualize the aggregated metadata
- To improve the quality of content and metadata across the BPN

The WP3 is in charge to organize the ingestion plan and the aggregation workflow from Content Providers to Europeana passing through EAGLE Aggregator; to monitor the work progresses and assure that the Tasks and Milestones will be satisfactorily achieved.

The Milestones are:

- Month 18 (September 2014): 375.000 items available for Europeana (MS8)
- Month 24 (March 2015): 750.000 items available for Europeana (MS12)
- Month 30 (September 2015): 1.125.000 items available for Europeana (MS15)
- Month 36 (March 2016): 1.500.000 items available for Europeana (MS16)

The Deliverable 3.2 'Ingestion plan' reports about the base of the further and future steps of the aggregation process and of the fundamental cooperation set up among the others Work Packages.

At this phase of the work, the cooperation between WP2 and WP4 is of paramount importance for the good planning of the structure that has to sustain the next phases of the project. The network and relations facilitating carried out by WP2, the content assessment monitored by WP3 (plus the development of the common EAGLE metadata schema), the tools for the aggregation developed by WP4 and the training sessions and material produced to inform the Providers about the technical requirements, help to build the ingestion workflow.

This document reports about the work carried out from the beginning of the project, month 1 (April 2013), to the month 12 (March 2014).

1.2 EAGLE AND THE EUROPEANA SISTERS PROJECTS

'The idea of a Cultural Commons is fundamental to the successful operation of a web ecology of content and services. Underpinning the foundation of this Commons is a set of resources in the public domain that are owned collectively or 'held in common' and shared openly among the community. With the release under a CC0 waiver of more than 20 million metadata records in their repository, Europeana has collectively taken a major step towards the goal of opening up data to enable access. Instead of trying to bring the user to Europeana, Europeana wants to take the material to the user. Europeana is doing this by developing strategic partnerships, by paving the way for creative re-use by developers and by providing the infrastructure that offers opportunities for creating new meaningful ways to access and interpret culture'¹.

The EAGLE Project (Europeana network of Ancient Greek and Latin Epigraphy) is a Best Practice Network² that is part of Europeana. The project is therefore inserted in this environment and will work within Europeana, and with its sister projects, to ensure full and effective integration within this flagship project to make European culture globally available.

The EAGLE consortium brings together institutions from previous Europeana sisters projects (just to cite one the Linked Heritage Project³ where the Cyprus Institute and La Sapienza University of Rome collaborated), partners with a long experience in digital initiatives (e.g. the University of Heidelberg, the University of Bari, the University of Salzburg and the University of Alcalá). There are also 'new entries' (e.g. Budapest, Cluj Napoca, Ljubljana, Pula Universities) that therefore need much more effort and assistance in harmonising their content or approaching the guidelines and requirements of the EAGLE project and of Europeana.

¹ Wikimedia source: Europeana/Projects. <http://meta.wikimedia.org/wiki/Europeana/Projects>. Accessed on 6 March 2014.

² Best practice networks (BPN) promote the adoption of standards and specifications for making European digital libraries more accessible and usable by combining the "consensus building and awareness raising" function of a thematic network with the large-scale implementation in real-life context of one or more concrete specifications or standards by its members. Each BPN tries out, on a sufficient mass of content, one or more of the implementation approaches discussed in the network in order to draw conclusions on their validity and if necessary to adapt them.

³ <http://www.linkedheritage.eu/>

As a further information, The Cyprus Institute (leader of the WP3) has a long experience in the coordination of content and assistance to the ingestion procedure for publishing content into Europeana (e.g. ATHENA Project, Linked Heritage project, AthenaPlus project)⁴.

Therefore, similarities in the setting up of the ingestion procedure may occur. Within EAGLE the presence of new partners in the Europeana environment, the use of different aggregations tools and metadata formats and the changes/updates of the Europeana guidelines (for example, the new metadata model developed by Europeana, EDM – Europeana Data Model and the new rights statements) determine new outputs and different solutions.

1.3 THE ROLE OF THE DELIVERABLE 3.2 IN THE PROJECT

The aim of the Deliverable 3.2 is the preparation and implementation of the ingestion plan. This is a fundamental step to assess and prepare the aggregation workflow: it is the preparation for the elaboration of the following deliverables and work. It is important for the relation with the further Deliverables planned during all the time-span of the project that will report on the effective contribution of the content to the EAGLE infrastructure and to Europeana by the Network Content Providers.

Since the beginning of the project (Kick-off meeting in Rome, 2-3 April 2013) a first draft of the general steps to be followed has been presented to the Content Providers. During the following months and meetings (First Plenary meeting in Pisa, 24-25 October 2013 and Second Plenary meeting in Ljubljana, 18 February 2014) more technical information were delivered according to the development of the parallel works of WP3 and WP4.

In fact, this is an important scenario to be shown to the partners to plan the work and the next Deliverables. The next Deliverables are strictly connected with the achievement of the ingestion procedure and of the Milestones. They are the MS8 (Month 18) with the ingestion of 375.000 items available for Europeana; MS12 (Month 24) with the ingestion of 750.000 items available for Europeana; MS15 (Month 30) with the ingestion of 1.125.000 items available for Europeana; MS16 (Month 36) with the ingestion of 1.500.000 items available for Europeana.

Table I gives an overview of the Deliverables and Milestones schedule:

MONTH	12	18	24	30	36
DELIVERABLE	D3.2	D3.3.2	D3.3.2	D3.3.3	D3.3.4
MILESTONE	-	MS8	MS12	MS15	MS16

Table I. The schedule of the WP3 Deliverables and Milestones. In light blue, the current report period.

The Deliverable 3.2 reports about the ingestion plan and the planning of delivery of EAGLE content to Europeana. This plan will help to foresee eventual bottlenecks, amendments and remedy actions that could be needed during the time of the project.

The ingestion plan and the work of the WP3 is connected to the work of WP4 for what concerns the full set up of the technical infrastructure and the preparation of the technical requirements useful for the ingestion of the content in Europeana. The ingestion plan is depending on the EAGLE metadata model (the metadata model and the related Deliverable, D3.1, have been delivered in December 2013 -Month 9 by UAH) and on the release of the AIM Infrastructure for the content ingestion by CNR (the first specifications and deliverable D.4.1 has been delivered in September 2013 – Month 6; the first release

⁴ Vassallo V., Piccininno M. (2012); Piccininno M., Vassallo V., (2013).



and related Deliverable D4.2.1 is due at March 2014 – Month 12, therefore at the same time of the present Deliverable). Starting from April 2014 the ingestion platform and the ingestion procedure will be active and it will be carried out the workflow already planned.

A first training and introduction to the ingestion infrastructure has been shown during the Second Plenary meeting in Ljubljana (18 February 2014).

Since the WP3 involves the content of the project and around the content are developed all the other activities, it is of great importance for the Work Package 3 tasks, to monitor also the progresses of the others WPs. This activity is carried out through the participation to the other WPs meetings and keeping update about their progress (see 3.5.4 in this Deliverable).

2. CONTENT PROVIDERS AND DIGITAL COLLECTIONS

1.4 CONTENT PROVIDERS

The EAGLE Consortium involves well-established research institutions and universities in the field of ancient epigraphy and cultural heritage. Some of them have already been involved in Europeana ecosystem projects (e.g. ECLAP, Linked Heritage, Athena).

The project Consortium, as described in the Description of Work, is composed of 19 partners from 12 different European countries. Within the EAGLE Consortium, 14 of them are Content Providers and they will supply Europeana with Ancient Greek and Latin inscriptions with related metadata: 1.5 M images and related metadata, (including translations of selected texts in Wikimedia) for the benefit of the public.

Table II shows the participating partners to the project, specifying the Content Providers (in light blue).

N.	Organisation	Country	Content Provider
1	UNIVERSITA DEGLI STUDI DI ROMA LA SAPIENZA	ITALY	✓
2	PARIS-LODRON-UNIVERSITÄT SALZBURG	AUSTRIA	✓
3	KATHOLIEKE UNIVERSITEIT LEUVEN	BELGIUM	
4	THE CYPRUS RESEARCH AND EDUCATIONAL FOUNDATION	CYPRUS	✓
5	EUREVA SAS	FRANCE	
6	CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE	FRANCE	✓
7	RUPRECHT-KARLS-UNIVERSITAET HEIDELBERG	GERMANY	✓
8	DEUTSCHES ARCHAOLOGISCHES INSTITUT	GERMANY	✓
9	EÖTVÖS LORÁND TUDOMÁNYEGYETEM	HUNGARY	✓
10	UNIVERSITA DEGLI STUDI DI BARI "ALDO MORO"	ITALY	✓
11	THE BRITISH SCHOOL AT ROME	ITALY	✓
12	PROMOTER SRL	ITALY	
13	GOGATE SRL	ITALY	
14	CONSIGLIO NAZIONALE DELLE RICERCHE	ITALY	

15	SVEUCILISTE JURJA DOBRILE U PULI	REPUBLIC OF CROATIA	✓
16	ZNANSTVENORAZISKOVALNI CENTER SLOVENSKE AKADEMIJE ZNANOSTI IN UMETNOSTI	REPUBLIC OF SLOVENIA	✓
17	UNIVERSITATEA BABES BOLYAI	ROMANIA	✓
18	UNIVERSIDAD DE ALCALA	SPAIN	✓
19	THE CHANCELLOR, MASTERS AND SCHOLARS OF THE UNIVERSITY OF OXFORD	UNITED KINGDOM	✓

Table II. List of the EAGLE Partners and Content Providers.

As already highlighted, beyond partners with a long experience in digital initiatives, many partners never contributed to Europeana and they will do it with a critical mass of quality content that is completely missing so far. They will bridge this gap within Europeana by making available a comprehensive collection of unique historical source, representing an important pillar of the European culture.

Furthermore, through the EAGLE consortium it will be possible to contribute content from countries that are less represented in Europeana, such as Romania, Croatia, and Cyprus.

1.4.1 The enlargement of the Consortium: new Content Providers

One of the aim of WP2 'Networking and best practices' is to attract new content providers to the EAGLE network, in order to increase the quantity and quality of the EAGLE digital resources and to ensure the wide array of scattered collections is covered appropriately. Content providers from outside the Consortium will be encouraged to join the Best Practice Network and to contribute their content. The partners who already expressed their interest and the intention to cooperate with the project will be the starting point for the enlargement of the network. This action will have as a project impact, a higher quantity of quality content available through Europeana. All the content providing partners fulfil an important role in awareness-raising and recruiting new content providers across Europe.

At March 2014, 4 affiliated partners signed the agreement to provide content to the project. They are: Pontificia Commissione di Archeologia Sacra, the Israel Museum, Jerusalem, Attic Inscription Online and the University of Beograd.

WP3 will assist WP2 to carry out these activities of consortium enlargement and to elaborate an action plan for the institutions that have shown interest in participating to the project activities. Particularly, WP3 will monitor the involvement of the stakeholders that will contribute with content to Europeana⁵.

⁵ For the methodology and the activities about the enlargement of the network, please refer to the Deliverable D.2.1 'Networking infrastructure and terms of reference'.

1.5 EAGLE COLLECTIONS

As declared in the Description of Work, EAGLE project will provide an important and vast collection of ancient writings on ancient artefacts to Europeana. At the moment, only a very small amount (about 30.000 items provided through the Linked Heritage project) has been provided to the European portal (Fig. 1).

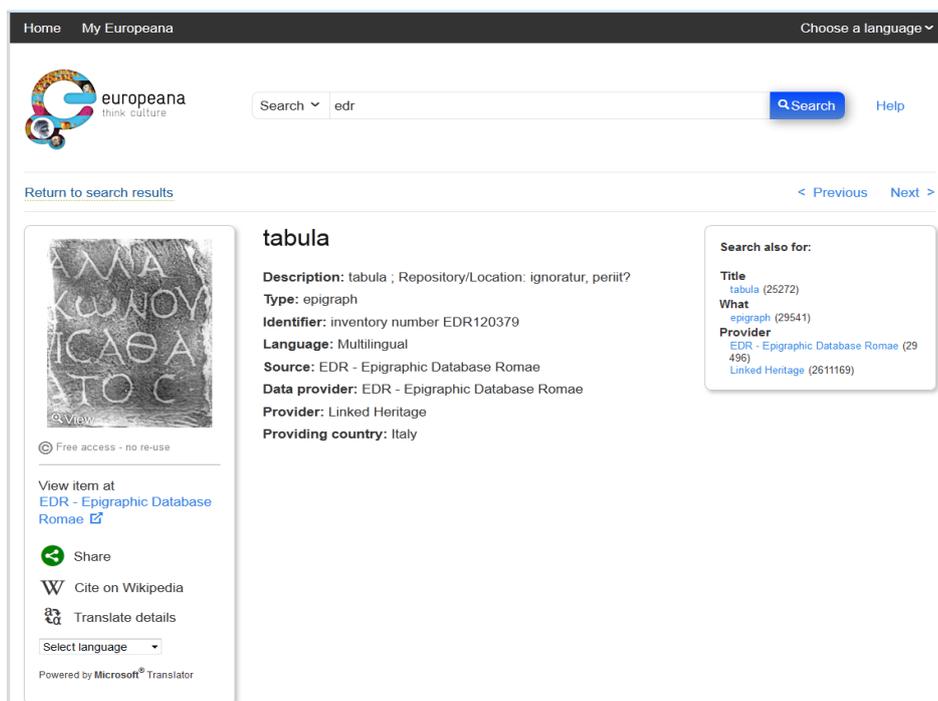


Fig. 1. Example from the small amount of digital inscriptions present in Europeana so far.

The content provided by EAGLE partners will allow users to take advantage of a high quality and unified repository consisting of all the inscribed material relevant to the ancient history of Europe and of the Mediterranean basin. In this way EAGLE will enrich Europeana with a great variety of inscriptions written in Greek, Latin and other ancient languages, providing scholars with an authoritative resource to verify the reliability of historical reconstructions, and the broad public with a way to understand interesting and curious inscriptions, geographically dispersed, only apparently difficult to appreciate. The ancient roots of the European culture will be made accessible. Inscribed monuments include authentic masterpieces of ancient architecture such as the Pantheon and the Trajan Column in Rome, the aqueduct of Segovia in Spain, and so on.

The EAGLE consortium will aggregate almost 1.5 million images and texts in digital form. This has to be considered a critical mass, since it comprises the vast majority of all inscribed material known at present, coming from the wide geographical area of the classical world, corresponding to a large portion of modern Europe, and including also the countries of the Southern and Eastern coasts of the Mediterranean basin from Morocco to Turkey.

The content is constituted of digital images (photographs, drawings, scanned notes, manuscripts and archive material, ancient books) and their related metadata.

Table III gives an overview description of the collections that will be provided by the EAGLE project:

UNIROMA1	EDR (Epigraphic Database Roma), an online, freely accessible, database (www.edredr.it), hosted by Digilab – Mediateca delle Scienze Umanistiche at Sapienza University of Rome, is going to provide all the about 60.000 Greek, Latin and bilingual inscriptions, from the VII century B.C. to the VII century C.E., discovered in the area of ancient Italy, including Sicily and Sardinia.
UNIBA	EDB (Epigraphic Database Bari), an online, freely accessible, database (http://www.edb.uniba.it), hosted by the University of Bari Aldo Moro, include Christian inscriptions of Rome until the end of the VII century C.E., recording texts in Greek, Latin, ancient Hebrew and old English.
UHEI	EDH (Epigraphische Datenbank Heidelberg), hosted by the University of Heidelberg and sponsored by the Heidelberg Academy of Sciences, includes three online, freely accessible databases: Epigraphic, Photographic and Bibliographic (http://www.uniheidelberg.de/institute/sonst/adw/edh/index.html). They all are focused on the Greek, Latin and bilingual inscriptions (including Punic texts) from the former provinces of the Roman Empire. UHEI will also provide photographs of Greek and Byzantine inscriptions of Aphrodisias made available by the King's College London.
UOXF	The Oxford University will provide the content of 'Last Statues of Antiquity', a catalogue of the Greek and Latin inscriptions for statues set up all over the empire from AD 284 onwards.
UAH	HE (Hispania Epigraphica Online), an online freely accessible database (http://www.eda-bea.es) hosted by the University of Alcalà de Henares, includes all the ancient (Greek, Latin and Palaeohispanic) inscriptions from the Roman provinces of the Iberic peninsula, corresponding to the modern countries of Portugal and Spain. With the cooperation of the Archivo Epigrafico de Hispania, digital items related to these epigraphic texts will be provided, including photos, file cards and scholarly notes, drawings from manuscripts.
PLUS	CHC (Research Group for Archaeometry and Cultural Heritage Computing), hosted by the University of Salzburg, will provide in total photographs of Roman stone monuments and inscriptions, including object photos of the UBI ERAT LUPA database (http://www.ubi-erat-lupa.org/platform_e.shtml), and scientific images concerning the archaeometric provenance of the objects.
UBB	The Babes-Bolyai University of Cluj Napoca will provide digital images of Greek and Roman inscriptions from the Roman province of Dacia and from Drobocea.
ELTE	The University of Budapest will contribute with images of the stone monuments found in the excavations of the Roman town of Brigetio.
UNIPU	The Juraj Dobrila University of Pula will provide 2000 images (photographs and drawings) of the Latin inscriptions from the ancient Roman towns of Colonia Iulia Pola and Colonia Iulia Parentium and their territory.
ZRC SAZU	ZRCEpigr. (Epigraphic Archives of Slovenia), hosted by the Institute of Archaeology of Ljubljana, will provide images (photographs, drawings and maps) of the Latin inscribed monuments from Emona and its territory.
AUSONIUS	The Institut Ausonius - CNRS - University of Bordeaux 3 will contribute with its archive of images of Greek and Latin inscriptions from France, Spain, Tunisia,

	Turkey, Russia, Ukraine.
DAI	The Deutsches Archäologisches Institut will contribute with its archive of images (photographs, itchings, drawings, scans of ancient books and the old volumes of the CIL) of Greek and Latin inscriptions, ancient monuments with inscriptions, sites where inscriptions were found from the whole Mediterranean and the North-Western provinces of the Roman Empire.
CYI	The Cyprus Institute will contribute with epigraphic inscriptions, translations, comments and images of ancient Cypriot literature and archaeology of the Archaia Kypriaki Grammateia Digital Corpus.
BSR	The British School at Rome will contribute with photographs of Greek and Latin inscriptions from Libya (Cyrenaica and Tripolitania) and South Etruria (Italy).

Table III. Description of the collections provided by EAGLE Content Providers.

3. THE INGESTION PLAN

The tasks of the WP3 started at the first month (April 2013) and will last for the whole duration of the project (36 months): indeed the works related to the WP started before the effective beginning of the project. Through meetings and communications, the leaders of the WPs involved in the ingestion process, agreed on the general procedure to be taken and the future steps.

Particularly, the attention was focused on specific information and documentation to be circulated among the partners in order to inform:

- how to provide content to Europeana (general aggregation procedure and general methodology);
- about the technical requirements needed;
- the assessment of the collections declared in the Description of Work by the partners.

All this information is in fact connected to the preparation of the activities and plans to arrive to the effective start of the Work Package (month 1) and to the completion of the project tasks (by month 36).

The ingestion plan summarizes all the activities on which is based the EAGLE aggregation procedure and the content delivery to Europeana: a detailed working plan composed of surveys, training sessions, partners support, procedures, methodologies, feedback, milestones. In detail, the WP3 aims at the planning and preparing the ingestion and at its implementation and monitoring.

For what concerns the first part (T3.3. Planning and preparing the ingestion), each content provider participate in the WP, carrying out work on selecting the content, harmonising the metadata, enriching and creating additional metadata if necessary, preparing the master object files for ingestion in the repository, preparing the surrogates, applying unique identifiers, etc. The task of the WP3 leader is to ensure that the Content Providers make their content available according to the project schedule; to ensure that the content providers provide the content declared; to ensure that the content providers have cleared the related IPR issues by the time they provide their content.

For what concerns the second part of the tasks (T3.4. Implementing and monitoring the ingestion), the Content Providers apply the best practices and implement locally the necessary workflows and procedures to ensure the continuous provision of metadata, surrogates and digital content through the EAGLE Aggregator (WP4) to Europeana. WP3 is in charge of monitoring the progress of the content ingestion into Europeana. The periodic report about this part of the work will start on month 18 (September 2014).

1.6 THE INGESTION PLAN FORM

In order to perform the above-mentioned Tasks, one of the most important action is the preparation of a working tool, in order to keep under control the delivery to Europeana of the content declared in the Description of Work: the ingestion plan form.

The ingestion plan form is useful in fact to keep trace of the Content Providers information about the items they will provide to Europeana, along all the project period. It is a useful tool to confirm or update the content declared in the EAGLE Description of Work and to trace the ingestion progress. To collect these information from the Content Providers, a control system document has been created, in order to describe the status and the availability of the collections to be ingested by each partner, starting from the table contained in the Description of Work (DoW) and monitoring continuously the providers' situation. This activity needs a constant update and revision of the content list: the results are gathered in a comprehensive table that gives a general overview of the situation, and partner by partner (Table IV).

Number	Short Name	DEA	DoW	Second Survey	MS8 (September 2014)
1	UNIROMA1	✓	60.000 images (photographs and drawings)	60.000	...
2
TOTAL			1.472.198		

Table IV. Example of the ingestion plan form: the table is a working tool for WP3 to keep track of the project progress.

1.6.1 The first Survey (March 2013)

Within the WP3 activities, the elaboration and launch of a survey was performed. The main aim of this work was to assess the information declared in the Description of Work (before the beginning of the project) in order to update the data and to prepare the future content ingestion plan. To gather all the information about the collections, Content Providers were requested to fill a questionnaire and to give examples of the metadata models used, in order to assess them for the creation of the new metadata model for EAGLE.

This first preliminary survey was distributed among the Content Providers some time before the effective beginning of the project (March 2013) in order to plan the future activities and to arrive to the start of the project with an organized plan. The survey, elaborated by WP3 in collaboration with WP4, was necessary to confirm or update the content declared by the partners in the EAGLE Description of Work (DoW). Moreover, the survey was useful to assess the technical requirements and specifications of the archives and digital libraries of the Providers.

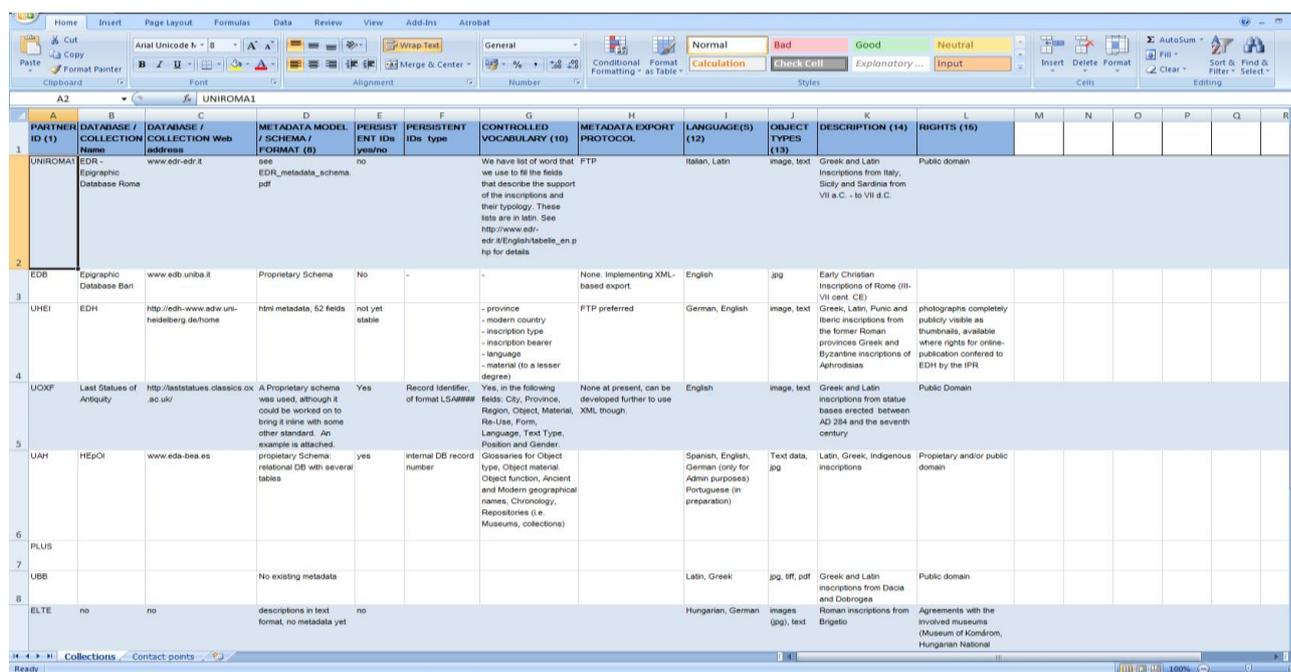
The table consists of the following fields and sub-fields, where the Content Providers were asked to give information about their collections:

1. Partner name (ID)
2. Database/Collection
 - a. Name
 - b. Web address
3. Metadata model/Schema/Format
4. Persistent identifiers
 - a. Yes/no
 - b. IDs type
5. Controlled vocabulary

6. Metadata export protocol
7. Languages
8. Object types
9. Description
10. Rights

The survey was launched at the beginning of the project by WP3 (Cyl) within the Task 3.3 in order to have a first update of the information declared by the Content Providers in the Description of Work and to confirm the content to be provided to Europeana. In this occasion, further information have been asked.

For the aims of the Task 3.1 it has been requested to the Content Providers the metadata schemas used in their archives and to provide an example of each one. The information and documentation gathered through the Survey have been put at disposal of all the WP3 Task Leaders and of the WP4 (Fig. 2). See Annex I.



A2	B	C	D	E	F	G	H	I	J	K	L	
PARTNER ID (1)	DATABASE / COLLECTION Name	DATABASE / COLLECTION Web address	METADATA MODEL / SCHEMA / FORMAT (8)	PERSISTENT IDs yes/no	PERSISTENT IDs type	CONTROLLED VOCABULARY (10)	METADATA EXPORT PROTOCOL	LANGUAGE(S) (12)	OBJECT TYPES (13)	DESCRIPTION (14)	RIGHTS (16)	
1	UNIROMA1	EDR - Epigraphic Database Roma	www.edr-edr.it	see EDR_metadata_schema.pdf	no	-	FTP	Italian, Latin	image, text	Greek and Latin inscriptions from Italy, Sicily and Sardinia from VII a.C. - to VII d.C.	Public domain	
2	EDB	Epigraphic Database Bari	www.edb.uniba.it	Proprietary Schema	No	-	None. Implementing XML-based export.	English	.jpg	Early Christian inscriptions of Rome (III-VII cent. CE)		
3	UHEI	EDH	http://edh-www.adw.uni-heidelberg.de/home	html metadata, 62 fields	not yet stable	-	FTP preferred	German, English	image, text	Greek, Latin, Punic and Italic inscriptions from the former Roman provinces Greek and Byzantine inscriptions of Aphrodisias	photographs completely publicly visible as thumbnails, available where rights for online-publication conferred to EDH by the IIR.	
4	UOXF	Last Statues of Antiquity	http://laststatues.classics.ox.ac.uk/	A Proprietary schema was used, although it could be worked on to bring it in line with some other standard. An example is attached.	Yes	Record Identifier, of format L54###	Yes, in the following fields: City, Province, Region, Object, Material, Re-Use, Form, Language, Text Type, Position and Gender.	None at present, can be developed further to use XML though.	English	image, text	Greek and Latin inscriptions from statue bases erected between AD 284 and the seventh century	Public Domain
5	UAH	HEP-01	www.eda-bea.es	proprietary Schema relational DB with several tables	yes	internal DB record number	Glossaries for Object type, Object material, Object function, Ancient and Modern geographical names, Chronology, Repositories (i.e. Museums, collections)	Spanish, English, German (only for Admin purposes) Portuguese (in preparation)	Text data, .jpg	Latin, Greek, Indigenous inscriptions	Proprietary and/or public domain	
6	PLUS											
7	UBB			No existing metadata				Latin, Greek	.jpg, tiff, pdf	Greek and Latin inscriptions from Dacia and Dobrogea	Public domain	
8	ELTE	no	no	descriptions in text format, no metadata yet	no			Hungarian, German	images (.jpg), text	Roman inscriptions from Brigetio	Agreements with the involved museums (Museum of Kónórdrom, Hungarian National	

Fig. 2. Results of the first survey organized in a global table in order to give an overview of the state of art at the beginning of the project.

After this phase, WP3 (UAH) requested a complete description of the metadata model the Content Providers were using at that time, including what is mandatory, what has multiple values, the detail on the types and vocabularies, etc. (see Deliverable 3.1 'EAGLE metadata model specification').

Table V resumes some important information gathered with the first Survey and used by WP3 and WP4 purposes:

Collection/Provider	Metadata model/schema/format	Export protocol
UNIROMA1	EDR metadata schema	ftp
EDB	Proprietary schema	-
UHEI	html metadata	-
UOXF	Proprietary schema	-
UAH	Proprietary schema	-
PLUS	Proprietary schema	-
UBB	-	-
ELTE	-	-
UNIPU	-	-
ZRC SAZU	-	-
AUSONIUS	TEI	OAI-PMH
DAI	TEI, METS+MODS, RAW XML data, CIDOC CRM, Dublin Core	OAI-PMH
CYI	Proprietary metadata schema	ftp
BSR	TEI	-

Table V. Main information gathered with the First Survey

1.6.2 The second Survey (January 2014)

The second survey was elaborate by WP3, with the support of WP2 and WP4, for the effective and practical ingestion procedure scope. The aim of the survey is in fact to review and assess the content to be aggregated within EAGLE and to retrieve further and updated information about the digital collections that Providers committed to give to the project, and data export.

The survey has launched at month 10 (January 2014), almost at the end of the first year of the project and before the effective ingestion procedure. The time has been chosen on the base of the fact that on December the EAGLE common metadata model was delivered and on month 12 (March 2014) is due the First Release of the ingestion infrastructure. Therefore, for the Providers this was the right moment, since they had a clearer knowledge of the technical requirements and they had already started to prepare their content according to the technical specifications requested (the common metadata model and the ingestion platform requirements) by the technical teams.

Only the partners that will provide Europeana with digital cultural content were requested to fill the form in. It was asked to give detailed information about the collections to be provided and if there are any variations respect to the Description of Work; or if they are encountering any delays or problems for the content provision.

The reason of this further assessment is due to the fact that, on the base of the past digital libraries projects experiences, some changes and variations in the content provision may occur. Moreover, WP3 decided to control the state of art of the collections, plus asking further questions connected with the requirements developed by the technical work packages and tasks, because of the 'novices' partners in the Europeana environment.

The second survey consists of two parts (see Annex II). The first one is dedicated to the Provider information: name of the organization, country, and contacts (it was asked to the providers to give direct contacts useful for speed communications, in order to clarify any doubts directly).

1. Provider information	
Information about the organization providing content:	
Name of the organisation	
Short name of the organisation	
Country	
Website	
Contact person for managerial questions:	
Name	
E-mail	
Telephone	
Skype ID	
Contact person for technical questions:	
Name	
E-mail	
Telephone	
Skype ID	
2. Content description	
Description of the collection(s) to be provided to EAGLE. Please duplicate the table for each collection, if needed.	
Collection name	Please, specify the collection as in the DoW
Collection URL	If the collection is not online yet, please specify the reasons (e.g. digital library under re-engineering, etc.) and the estimated date of publication

Amount of digital objects	Amount of digital objects linked to the metadata. NB: The ratio between metadata and digital objects (1 d.o : 1 metadata; n d.o : 1 metadata)
Amount of metadata	Amount of metadata that you will aggregate within EAGLE. NB: Europeana aggregates metadata linking to digital objects
Object type(s)	Object types according to Europeana: IMAGE, TEXT, SOUND, VIDEO, 3D
Metadata export format	Please indicate whether your metadata records already comply with EpiDoc as agreed and described in Deliverable 3.1. If not, please state an estimation about when the compliance will be met. Cyprus, Lupa and images from Arachne will be collected in their own original metadata format (as they do not fit in EpiDoc properties) and mapped to Entities/properties identified in the EAGLE Aggregation Metadata Schema.
Protocol to export metadata	Please specify by which means (XML records provided via OAI-PMH, FTP, HTTP) are you intended to expose/offer your metadata to the EAGLE aggregation system. For those providers passing through EDH and EDR, please specify it. NB. More about this topic will be covered during the training session in Ljubljana, so just state your possibilities.
Amount of metadata for the first testing	In order to test the EAGLE aggregation platform, it is recommended that all the partners provide a sample of their metadata. Please specify the amount and the sample you provide for the ingestion platform testing.
Available rights	Please specify the rights that are on the digital objects. Europeana requires that data providers apply a rights statement for all digital objects described in their metadata. The rights that apply to the digital object will also apply to the previews used in the Europeana portal. The available rights statements are published at this page http://pro.europeana.eu/web/guest/available-rights-statements . Please specify if you still need to clarify your rights.
Notes	Please add any further information you think is useful.

Fig. 3. Example of the second Survey: the fields about the content description.

The second part (Fig. 3) is dedicated to the content, and it is structured as follows:

- Collection name
 In this field the Providers have to provide the exact name of the collection and specify to which one corresponds in the DoW.
- Collection URL
 In this section they have to clarify if the collection is online and, if not, the reasons (e.g. digital library under re-engineering) and the estimated date of publication.
- Amount of digital objects
 Providers were asked to confirm the amount of digital objects and to declare the ratio with related metadata (Example. 1 digital object : 1 metadata; n digital objects : 1 metadata)
- Amount of metadata
 In this field, it is asked the amount of metadata that Providers will aggregate within EAGLE, in reference to the digital objects declared in the Description of Work.
- Object types
 Providers were asked to specify which object types, according to Europeana requirements, will be supplied to the Europeana portal: IMAGE, TEXT, SOUND, VIDEO, 3D.
- Metadata export format

In this field, it is asked to indicate whether metadata records already comply with EpiDoc, as agreed and described in Deliverable 3.1. If not, it is asked to estimate about when the compliance will be met⁶.

- Protocol to export metadata

It is asked to Providers to specify by which means (XML records provided via OAI-PMH, FTP, HTTP) are intended to expose/offer their metadata to the EAGLE aggregation system. Since some providers will be aggregated through other digital libraries (in particular through EDH and EDR), it is asked to the Providers to specify that⁷.

- Amount of metadata for the testing phase

In order to test the EAGLE aggregation platform, it is recommended that all the partners provide a sample of their metadata. For this reason in this field it has been asked to specify the amount and the sample they provide for the ingestion platform testing.

- Available rights

In order to clarify the rights on the digital objects (metadata are under CC0 licence), it has been asked the providers to specify them. Europeana requires that data providers apply a rights statement for all digital objects described in their metadata. The rights that apply to the digital object will also apply to the previews used in the Europeana portal⁸.

- Notes

In this field Providers are asked to declare any further information they think is useful for the aim of the ingestion and that may cause any problems to the ingestion procedure.

One of the aim of WP3 is to coordinate carefully the content provision to Europeana through the EAGLE infrastructure. Therefore, WP3 operates a continuous monitoring about the content amount to be provided and checking of any possible delays or problems that may have a negative impact on the content provision. For this reason, Content Providers are also asked to specify the amount of both metadata and digital objects, in order for the WP3 and WP4 Leaders to verify the digital objects amount declared in the DoW, and to make sure they are the same that will be provided to Europeana.

For what concerns the results of the second survey, all Content Providers already confirmed the content declared in the Description of Work, apart from some exceptions: EDB, BSR, Ubi Erat Lupa, ZRC SAZU, Petrae. This difference is due to different reasons that are under monitoring and that are highlighted in the table below (permissions to publish the images, over estimation of the collection at the time of the DoW, availability of the images, signature of the DEA)⁹. In any case, also these Content Providers confirmed

⁶ Cypriot content, Ubi erat Lupa and images from Arachne will be collected in their own original metadata format (as they do not fit in EpiDoc properties) and mapped to Entities/properties identified in the EAGLE Aggregation Metadata Schema.

⁷ During the last Plenary meeting in Ljubljana (18 February 2014), during the technical workshop, for those providers that do not have a specific exporting protocol, some possible alternatives were proposed (the provider will have the possibility to deposit the content in a FTP server set up by CNR and where to harvest the data from)

⁸ The available rights statements are published at <http://pro.europeana.eu/web/guest/available-rights-statements>.

⁹ Indeed, the work of the WP3 is that one to keep direct contact with each provider. Thanks to these periodic communications (see paragraph 3.5.1 of this Deliverable) it was possible to highlight the reasons of the temporary changes/delays and of the relations between digital objects and metadata (e.g.

that they would be able to cover the amount declared in the Description of Work by the end of the project, overcoming the problems encountered so far by replacing some collection or by involving new cultural institutions.

The Table VI gives an overview of the current changes of the items declared in the last survey respect to the amount declared in the Description of Work.

Number	Provider	DoW	Survey (January 2014)	Reasons/comments
1	UNIROMA1	60.000	60.000	
2	UNIBA	45.000	34.329	Permissions to publish the images
3	UHEI	65.000	65.000	
4	UOXF	1.500	1.500	
5	UAH	235.000	235.000	
6	PLUS	51.000	35.805	Over estimation of the collection at the time of the DoW
7	UBB	4.000	4.000	
8	ELTE	1.500	1.500	
9	UNIPU	2.000	2.000	
10	ZRC SALZU	400	300	Availability of the images
11	AUSONIUS	5.000	706	Signature of the DEA
12	DAI	1.000.000	1.000.000	
13	CYI	442	442	
14	BSR	1.356	1.342	Permissions for the images
TOTAL		1.472.198	1.441.924	

Table VI. The table shows the current changes in the content provision.

WP3 will constantly monitor the activities of the Content Providers to assure the content will be provided to EAGLE by the end of the project and that the amount declared will be covered. Further surveys and periodic interviews, together with the already planned and ongoing monitoring provider by provider, can be performed to observe the general content provision.

1.7 DATA EXCHANGE AGREEMENT (DEA)

Within the WP3 tasks, an important requirement for the content provision monitoring is the sign of the Europeana Data Exchange Agreement by all the EAGLE Content Providers. Indeed, as a Note reported in the EAGLE DoW, 'All beneficiaries submitting metadata to Europeana shall do so pursuant to the

enrichment and update of the metadata, publication online of the content, set-up of the technical requirements of the project, metadata description associated to multiple digital objects).

Europeana Data Exchange Agreement (DEA). Metadata submitted to Europeana shall be released as Open Data in accordance with the terms of the DEA. This requires that the IPR status of the digital objects thus made accessible is described by appropriate rights labels attached to them to allow users to know what uses they are allowed to make of the object¹⁰.

The subscription of the DEA it allows that metadata descriptions are subjected to the CC0 licence¹¹: the content is released as public domain¹² and this allows the commercial reuse of metadata and the possibility for Europeana to publish the metadata as Linked Open Data (LOD).

For this important reason, WP3 and WP1 (respectively for content monitoring and administrative control) keep under control the status of the DEA signature within the EAGLE consortium. Every Content Provider, supplying content to Europeana, has to subscribe the Agreement and therefore the WP3 task is to verify who signed and who is missing in order to keep under control the publication process in Europeana from the ingestion point of view.

At February 2014, only one Content Provider (Ausonius, France) did not sign the DEA yet, because they needed more time to verify and clarify the effects of the signature on their content. Since, like in this case, these providers never provided Europeana with metadata, WP1 and WP3 had to give the necessary instructions and information about it. At the moment, the Partner is studying the contract but they declared the intention to sign the agreement in order to fulfil their provision tasks.

1.7.1 Europeana Rights Statements

As previously explained, Europeana's Data Exchange Agreement requires that data providers apply a rights statement for all digital objects described in their metadata. The rights that apply to the digital object will also apply to the previews used in the Europeana portal. This rights statement is stored in the 'edm:rights' field of the Europeana Data Model (EDM). Each digital resource provided to Europeana has therefore supply the applicable right in order to inform the users about the terms under which the digital object and the corresponding preview can be used. Europeana has compiled a list of statements that can be applied to a particular object.

For this reason, since the beginning of the EAGLE project, WP3 informed the partners about the availability and the necessity to use these particular rights statements associated to the digital objects to be provided to Europeana.

On occasion of the second Survey (January 2014), WP3 asked to the Content Providers to list all the Europeana Rights Statement that will be used for their digital objects. The results of the survey, put at disposal of the WP2 for the research about IPR issues, highlighted the use of the following statements¹³:

- Rights Reserved-Free Access
- Creative Commons - Attribution, Non-Commercial
- Creative Commons - Attribution, Non-Commercial, ShareAlike

¹⁰ Such objects shall be labelled in accordance with the 'Europeana Rights Guidelines' available at <http://pro.europeana.eu/web/guest/technical-requirements>'.

¹¹ The digital objects are under the copyrights stated by the Content Providers according to the Europeana rights guidelines.

¹² <http://wiki.creativecommons.org/CC0>.

¹³ In few cases the Providers declared that have still to better clarify the type of rights to be used.

- Creative Commons - Attribution
- Creative Commons - Attribution, ShareAlike
- Unknown
- Public Domain CC0
- Creative Common – Attribution, No Derivatives
- Creative Commons - Attribution, Non-Commercial, No Derivatives

However, on February 2014, Europeana informed about the changes on their rights statements: in particular about the addition of new rights statements (e.g. "Orphan work" and "Out-of-Copyright - Non Commercial Use"), the change in the name and wording of a few statements and the removal of other (e.g. "Rights reserved - restricted access").

During the last Plenary meeting in Ljubljana (February 2014) and through the detailed communications circulated among the Consortium just after all, the Content Providers have been informed of the changes. In the next future, further interviews of the providers will be performed by WP3 and WP2, in order to clarify the current situation according to the new statements published by Europeana¹⁴.

1.8 CONTENT PROVIDERS PLANNING FOR THE DELIVERY OF EAGLE CONTENT TO EUROPEANA

The core of the WP3 concerns the provision of quality metadata and content from the individual partners of the Consortium, through the EAGLE Aggregation and Image Management (AIM) infrastructure implemented by WP4, to Europeana. The common EAGLE metadata, developed within WP3, is the basis for the local mappings to ensure the full interoperability with Europeana. Within WP3, different activities are addressed for the need to locally harmonise the content coming from different sources and prepare it for the ingestion process: the implementation of the agreed standards and protocols, the preparation of the necessary mappings for the metadata conversions, the actual ingestion and publication of the content to Europeana.

In this perspective, all the Content Providers actively participate in the planning of the delivery of EAGLE content to Europeana, carrying out internally all the activities oriented to perform that aim (Task 3.3).

Following, some reports of the work planned and carried out by the Partners in order to prepare and ingest their collections to Europeana through the EAGLE infrastructure.

UNIROMA1 – Epigraphic Database Roma (EDR)

“EDR will contribute to the EAGLE project with 60.000 digital objects and their related metadata. Uniroma1 will participate to the first testing ingestion phase with a dataset of 1000 items. Our data are structured in MySQL tables and are extracted from the Database and then exported in XML format, according to the EpiDoc schema requirements. The procedure to export in this format is implemented and the test with the first 1000 is already done. The remaining 59.000 items (metadata and digital objects) will be sent in the next months and by the end of the project. Our unique identifier is the id_nr, the internal identifier for the Database. The Trimegistos number will be applied directly on the EAGLE’ s server” (Silvia Evangelisti for UNIROMA1).

¹⁴ Europeana informed that by July 2014 the new statements have to be effective.

UNIBA – Epigraphic Database Bari (EDB)

“Enriching metadata. EDB is working on the enrichment of its records to reach the amount stated in the Description of Work. Since the beginning of project over 4.000 new records have been updated: 3.201 in process by compilers, 740 pending approval, 142 approved. The large number of records under completion is due to necessary update to recent changes applied to the database about inserting structured bibliographic references.

Harmonising the metadata. EDB contributed in this first year of the project to the creation of controlled vocabularies, especially on Decoration and Type of inscription. Consequently, EDB has contributed to the creation of tables of conceptual equivalence between the definitions in EDB and controlled vocabularies established for the EAGLE. This was necessary in particular for Material and Object type vocabularies, since the two classes are distinguished in the metadata model of EAGLE, but are not separated in EDB structure. Redundancies in the definition of executing techniques has been corrected by proposing the necessary changes to the metadata model that provides the indication of a single execution technique and a single conservation place, while the documents in EDB require more complex information.

Preparation of digital objects and of the surrogates. Regarding the "surrogates", they are necessary for lost inscriptions, which now correspond in EDB to slightly less than one-tenth of the total so far.

Regarding the digital objects of the inscriptions still existing, the agreement with PCAS allows us to count at the present state on 2330 Jpg files through an unique link between EDB and Archives PCAS. It is providing to the identification of ca 12000 photos available in the archive PCAS and other ca 8000 in the archive of EDB too. However, it will be impossible to obtain a digital photographic reproduction for each EDB record (there is no complete coverage for the Roman catacombs even by those directly responsible for the monuments, such as the PCAS, even due to insurmountable difficulties of shooting). Consequently, it is expected to adopt for these records the same solution "surrogate" required for lost items.

Applying of unique identifiers. Geographical data, originally in EDB simple free texts, are now traced to Geonames ID, as well as Repositories where possible were traced to URI established in Trismegistos.

IPR issues. About the images of inscriptions pertaining to monuments under the protection of PCAS, the issue of reproductive rights is overcome by the agreement between PCAS and EAGLE. About images of inscriptions pertaining to monuments under protection of other Vatican Institutions (Vatican Museum, Papal Basilicas), there is still no agreement. The only possible way is to forward a formal request from the European Consortium EAGLE to the competent authorities. For all the other cases EDB benefits from the agreement between Italian Ministry of Cultural Heritage and the previous EAGLE consortium.

Preparing the master object files for ingestion in the repository. A selection of 500 inscriptions will be provided for the first testing phase. Currently, EDB inscriptions are stored in a relational PostgreSQL database, but the procedure for exporting data to the EAGLE metadata model is under completion. In this respect, we will export data directly from the relational PostgreSQL database to the EAGLE metadata model, taking into account the vocabularies already been defined” (Anita Rocco for UNIBA).

UHEI - Epigraphische Datenbank Heidelberg (EDH)

“Selecting the content. Content curation at EDH continues at a smooth pace as usual, content will all be submitted, no selection required.

Format exporting. As by Deliverable 3.1, we export xml EpiDoc of inscription's text; we export xml in CIDOC-CRM for photos.

Set up of an export method: we currently run a server operation and call XSLT to export into a downloadable folder. FTP is foreseen.

Format and data currently available. XML is currently available for all inscriptions.

Schedule of provision. We have already provided data. We shall wait to add TM ids and TM geo ids to export the all data and give it for ingestion. XSLT are being prepared to populate the EpiDoc xml with tabular information provided.

Tasks being carried out internally. Such as enriching or creating metadata for collections, this continues to be done at EDH independently from the provision of data to the project. Photos and text are ready and available.

Applying of unique identifiers. URI are available and stable for all inscriptions.

IPR statements for photos are available for each photo". (Pietro Liuzzo for UHEI)

University of Salzburg (PLUS) – Ubi Erat Lupa

"We set up an xml export tool for the LUPA data: detailed info see <http://www.ubi-erat-lupa.org/eagle/>. The tool currently exports all the records from the database, but it does not yet include all the fields.

Currently, there is the availability of metadata on 22.318 physical objects and 35.814 related images, including appropriate copyright statements.

To reach the proposed No. of 51.000 images (DoW), which was an extrapolation done before the beginning of the project, could be reached in the course of the EAGLE project and the following activities are on schedule. We negotiate to integrate the objects from the Carnuntum Deposit (Austria), the Trier Museum (Germany) as well as the Split Musuem in Kroatia, which will bring ca. 4.000-5.000 objects and about the triple of the images. Beside that work, smaller contributions by various content providers are permanently running. The 1.000 petrographic images which are mentioned in the DoW are stored in the simplified petrography (<http://chc.sbg.ac.at/sri/thesaurus/>) which is the base for the material thesaurus in the EAGLE Project". (Christian Uhler for PLUS)

ZRC SAZU - Epigraphic Archives of Slovenia

"We currently sent to the database EDR 105 inscriptions and 53 illustrations, mainly photographs. We began submitting material by starting with the best published inscribed monuments. The export format is word document and jpg. All other criteria depend from the EDR, within which our material is being collected. Our next step will be to send to the EDR those epigraphic monuments that have not yet been published properly, some of which not having been collected yet. The tasks being carried out to achieve this scope are diverse and manifold. Early 16th century manuscripts of Augustinus Tyfernus, kept in the National Library in Vienna, must be revised, as well as other manuscripts and early publications of these inscriptions in various libraries. Autopsy should be made on the spot, which is often hindered by the inaccessibility of the monuments. These inscriptions must then be compared to their publication in the CIL III – if existing – and correctly identified. All this work is currently being carried out. As for the IPR issues, we came so far to a mutually satisfactory agreement with the National Museum of Slovenia, which possesses photographs of their epigraphic collection of ca. 100 monuments. We are also allowed (in terms of the common Europeana project), to use as illustrations the transcriptions of the inscriptions from Emona and its territory contained in the manuscripts of Augustinus Tyfernus. This is particularly important, since some of the monuments are lost. For the storytelling app. we submitted five stories". (Marjeta Šašel Kos and Anja Ragolič, for ZRC SAZU)

The Cyprus Institute - Archaia Kypriaki Grammateia Digital Corpus (AKGDC)

“The Cyprus Institute, beside the role of the WP3 leader, is also Content Provider to EAGLE project. The work consists of providing through the tool developed by WP4 the amount of metadata declared in the DoW. Within this task, Cyl has elaborated a new metadata schema for the cross disciplinary description of Cypriot inscriptions. The cross-disciplinary metadata for Ancient Cypriot inscriptions integrates information regarding the objects themselves and their digital “surrogates” (3D models, photographic documentation, digital texts, transliterations, etc.)¹⁵.

Cyl, as Content Provider, will afford Europeana the metadata of the Cypriot inscription corpus, enabling a wider visibility and increased awareness to the Archaia Kypriaki Grammateia Digital Corpus Project (AKGDC). This aggregation process will be possible through the mapping from the proprietary metadata schema to the EAGLE common metadata format developed within the project community and based on standards and recommendations and the service platform for epigraphy to allow for multi-format ingestion and multi-device. The declared content will be ingested in batches along the project time, starting with some digital objects and related metadata (approximately 0.4 % of the total amount) for the first testing phase and following progressively”. (Valentina Vassallo, for The Cyprus institute)

Juraj Dobrila University of Pula (UNIPULA) – Inscriptions of Pula

At the University of Pula we started to ingest inscriptions in EDR in November 2013. In front of us was the amount of 2,000 epigraphic monuments to be ingested in the database and until the end of February we ingested around 550 of them. The main centres where the epigraphic material comes from, are the Roman colonies situated in the peninsula of Istria: Pula and Poreč (Pola, Parentium) and their agri. First we started ingesting inscriptions from Poreč because their number is limited so we decided that it is better to do one corpus of epigraphic monuments in its integrity and to have in this way one part of the work completely done with all metadata and digital objects which it is possible to gather. This corpus of inscriptions from Poreč is now more less finished (just a few inscriptions are missing) and is available to the public through EDR. For this collection of epigraphic monuments we made photos in Poreč city museum Zavičajni muzej Poreč / Museo del territorio parentino (as much as it was possible, because some of them are lost) which gave us the permission to access inscriptions, take photos of them and publish it on the web. It is all now visible in EDR in the way that every photo took there has the mark ZMP/MTP Poreč/Parenzo and in information it is specified per gentile concessione ZMP/MTP Porec/Parenzo. Part of the epigraphic material from Poreč (Early Christian mosaics and graffiti) is related with Basilica Euphrasiana and located there mostly in situ. We got permission from the Diocese to make photos of them and publish it online, so they are now available with photos in EDR as well. Part of the inscriptions which are lost or unavailable to see, we scanned and ingested the photos or drawings mostly from Inscriptiones Italiae. Until now in EDR there are 290 inscriptions from Poreč and ager parentinus, 108 of them are completed with digital object (92 with photo of monument, 16 with scanned drawings). We are currently working on the epigraphic material from Pula (Pola) and we ingested until the end of February around 300 inscriptions from there all published in Inscriptiones Italiae. They are continuously appearing online how we ingest them (Valentina Zović for University of Pula)

1.9 WP3 PLANNING FOR THE CONTENT PROVISION

Work Package 3 is in charge of observing the content provision from the EAGLE partners to Europeana and of monitoring possible changes or delay, that might occur, controlling that the content will be

¹⁵ V. Vassallo, E. Christophorou, S. Hermon, F. Niccolucci, *Revealing cross-disciplinary information through formal knowledge representation – a proposed Metadata for ancient Cypriot inscriptions*. Conference Proceedings of Digital Heritage 2013, 28 Oct – 1 Nov 2013, Marseille.

maintained unchanged as declared in the Description of Work or as updated by the providers during the assessment.

At the time of the current Deliverable release, it is possible to report that the Providers almost confirmed the amount committed in the DoW and the contribution will be performed increasing the amount during all the project time-span.

As reported previously in the paragraph 3.1.2 about the results of the second Survey, there are no significant changes in the general content provision and the partners are working in overcoming the issues encountered in order to arrive to the Milestone 8 (September 2014) with the amount of 375.000 items aggregated in EAGLE and ready to be harvested by Europeanana.

WP3 asked all the Content Providers to contribute to the September deadline, distributing their work during the upcoming 6 months (April-September 2014). Therefore, all the Content Providers are currently working to meet the deadline and sending the content to WP4 in order to upload and map/transform their data in the technical infrastructure. The strategy chosen of making all the Providers to contribute, is made with the purpose to avoid any postponing by the Providers and force them to contribute equally and well distributed during the project time, trying to avoid “last minute” contribution or delays.

1.9.1 Mid-term internal milestones

On occasion of the second Survey, it has been asked the Providers to declare the amount of data they will provide for testing them in the technical infrastructure developed by WP4. By the end of the current month (March 2014), when it is also due the first release of the EAGLE infrastructure, the partners have to send the declared amount to the WP4 leaders in order to be tested¹⁶. Actually, this first batch of content for test is an effective contribution. The Table VII shows the amount of data declared by the Content Providers and that will be available on the EAGLE platform at the beginning of April 2014.

The strategy of the internal intermediate deadlines is aimed at a continuous monitoring and assessment, in order to check the work done and plan the next steps of the remaining period until the official milestones. Furthermore, the comparison of the results achieved, steps by steps, with the objectives and milestones foreseen, it is of great importance for the evaluation of possible bottlenecks and for the possibility to prevent and search for the right solution on time.

For this aim, it has been elaborated a working tool to be circulated among the Providers at predetermined intervals and according to the official Milestones. It is an excel table where the Content Providers have to declare the amount of data (a breakdown of the provision) they will deliver to the EAGLE infrastructure in order to be harvested by Europeanana.

In the Annex III, WP3 has performed a general forecast of the provision: on the base of the amount that Content Providers declared for the first internal deadline and in the DoW, an estimation of the percentage for each Provider is made, in order to calculate the effective provision and monitor the eventual deviation in the foreseen amount. A comparison between the estimated amount and what will be effectively ingested will be performed during the following months, in order to calculate the eventual differences and control the deadlines and tasks to be carried out by the providers.

The first round of content is going to be send at the end of the month (March 2014) and it will cover the 6 months from April to September 2014, in order to monitor the distribution of the Providers work during this period and catalyse the efforts of all the parts in reaching the project deadlines. The structure of the working tool, with the names of the providers, the amount to be declared by each one and the amount to be reached for the Milestone helps to monitor the work progress.

¹⁶ Some Providers already send a number of items for the infrastructure test aims, before its release.

PARTNER SHORT NAME	ITEMS FOR APRIL 2014
UNIROMA1	1.000
UNIBA	500
UHEI	1.000
UOXF	1
UAH	200
PLUS	35.805
UBB	800
ELTE	800
UNIPU	500
ZRC SAZU	108
AUSONIUS	10
DAI	50.000
CYI	10
BSR	1.016
TOTAL	91.750

Table VII. Amount of data that Providers are sending for the April deadline.

Since the amount of the Milestone 8 corresponds to the 25% of the total amount declared in the Description of Work, this means that each Content Provider should supply with the 25% of its content for September 2014. For this reason, WP3 asked to the partners to make an estimation of the provision during the next 6 months in order to cover the amount foreseen for the deadline¹⁷.

Furthermore, since also the next Milestones are distributed every 6 months, this method is the best solution to have a micro and holistic view of the Providers work during the entire project time span.

1.10 MONITORING AND SUPPORTING THE PROVIDERS

Implementing the ingestion plan in a digital library project means not only to monitor the work of the Content Providers in supplying content, but also to support them in achieving this goal. This determines a series of actions and activities able to establish a direct and fruitful relation with the involved parts.

1.10.1 Direct contacts

Since the beginning of the project, WP3 started a direct relation with the Providers. In the first survey, it was asked in fact to provide, apart from e-mail accounts for more formal communications, also direct contacts, such as Skype and telephone number, to establish a more direct and informal exchange of

¹⁷ Considering a smooth provision during the period April-September 2014, each content Provider should contribute monthly with about the 4% of its declared amount.

information. Therefore, WP3 is periodically and constantly in contact with the Content Providers through e-mails and Skype calls in order to monitor the progress of their work, to inform about news and updates or discuss about any kind of issues that they may encounter, clarifications of doubts, etc. These kind of contacts effectively substitute the valuable face-to-face meetings, usually carried out during the plenary meetings of the project.

In addition, WP3 prepared a working tool to keep track of all the information between the parts. Thanks to this simple procedure, it is possible to collect the Providers' feedback, to monitor their activities and check the ingestion process on a structured document. The monitoring plan follows the Content provider ingestion from the beginning of the project (uploading of the content in the EAGLE infrastructure) until the end (the publication in Europeana) and it is fundamental to monitor the available content and the progress of the project (Table IX).

This is a working document that helps to maintain updated the information and the communication the topics exchanged between WP3 and the Content Providers, drawing its personal history in a way fast to read and to consult. In detail, the sections gather the most important information related to the ingestion plan and to the provision of the content: what are the criticisms encountered by the Content Providers and can cause issues in the data provision. If there are any changes for the content to be delivered; the effective provision respect to the mid-term internal deadlines and the project Milestones, etc. Moreover, all the topics discussed are summarized and registered in the file with the date. As a matter of fact, the possibility to retrieve easily the communication and the issues, helps to make the process smoother and to identify on time any possible problems.

PROVIDER NAME
CONTENT SURVEY To check if the provider has updated the initial ingestion plan form supplied with the Description of Work (DoW) To keep track of the information given in the assessment
CHANGES RESPECT TO THE DoW Any change made to the DoW in terms of updates of the collections, amount of the collections, removal of collections and eventual substitutions.
DEA Did the provider sign the DEA?
DIGITAL OBJECTS = METADATA Ratio between the amount of digital objects and metadata
MILESTONES AND MID-TERM INTERNAL MILESTONES How many records have been provided each month? How many records have been provided for MS?
CRITICISMS List of the problems encountered by the Content Provider
COMMUNICATIONS Recap of the communications between WP3 and the responsible of the collections.

Table IX. Monitoring plan of the Content Provider

1.10.2 Training sessions and demonstrations

During the last two EAGLE Plenary meetings (24-25 October 2013, Pisa; 18 February 2014, Ljubljana) WP3 and WP4 organized practical sessions in order to train the Content Providers about the EAGLE metadata schema, on the use of the ingestion platform and about the ingestion procedure and aggregation to Europeana.

The idea of these training sessions is to keep informed and trained the Content Providers about the technical progress of the project, e.g. updated information about the EAGLE common metadata format and the ingestion platform, the modality of the mappings and delivery of the content to Europeana.

The first training session in Pisa was addressed to:

- the presentation of the EAGLE metadata format and its specifications (under finalization at that time and delivered at December 2013);
- to start training the Content Providers on the use of the schema, the content preparation, the mapping and uploading through the EAGLE ingestion tool (due in March 2014).

The second training session in Ljubljana was dedicated to:

- the training of the Content Providers about the mapping and tools developed for EAGLE. In detail, the technical part about the aggregation infrastructure, metadata preparation and export, the metadata mapping, the data flow and the transformation results;
- the procedure and deadlines of the content provision. In particular, the Content Providers were informed about the aggregation workflow, the ingestion plan and the system of monitoring their work, the publication on Europeana portal, the feedback gathering (from the Providers and from Europeana) and all the procedure deadlines.

This is part of the support of the Content Providers within the activities of the project (Fig. 4).



Fig. 4. The first training session and demonstration during the Plenary meeting in Pisa (October 2013).

The training sessions/demonstrations can be organized and held again during the time of the project (e.g. during the next Plenary meetings) and according to the needs of the Content Providers, in order to support their work and to solve any difficulties in the provision procedure to Europeana.

All the material about the training sessions, the presentations and useful documentation is put at disposal of the Content Providers and of the project partners on the EAGLE project repository (Fig. 5). The common space is maintained updated and the partners are asked to keep checking it.

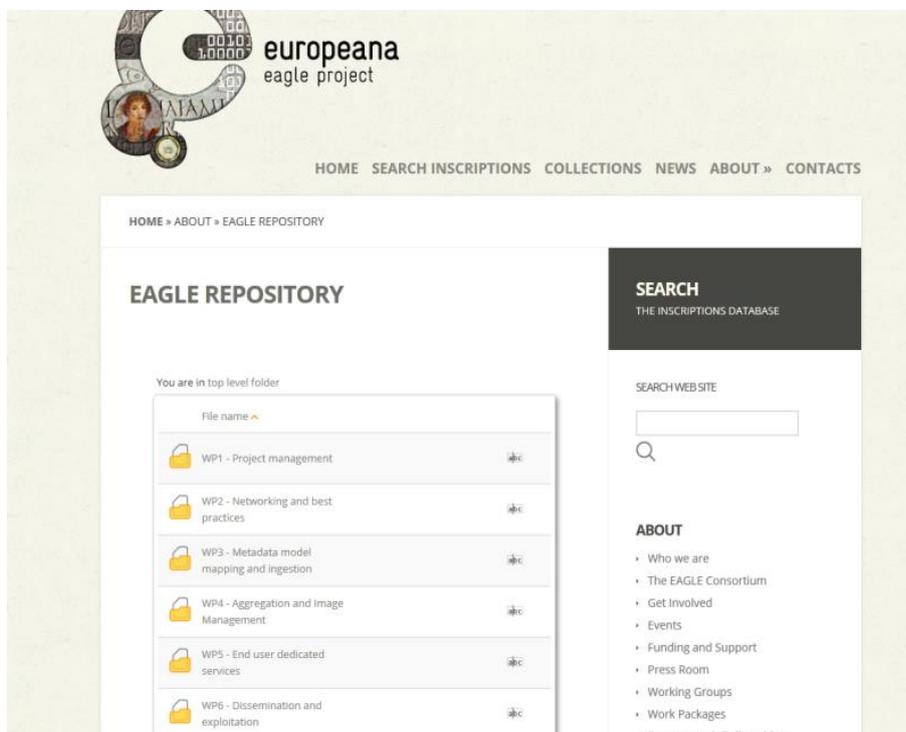


Fig. 5. The documentation of the training sessions is kept updated on the EAGLE Repository.

1.10.3 Setting-up of a Support Forum

During the Plenary and technical meetings in Pisa (24-25 October 2013) in agreement with the WPs leaders, the Content Providers and partners, it was decided to set up a technical support for the ingestion activities of the project.

The experience of several digital libraries projects within the Europeana environment, highlighted the importance to use a support tool able, not only to monitor the work of the Content Providers, but also to receive feedback, to detect and identify the possible issues and to keep trace of the work progress, both in general and in detail for each provider.

Particularly, the Consortium decided to choose the formula of the forum in order to facilitate the communication among the Providers and the partners as a common space for discussion, and to maintain open access for an easy finding of the threads topics. This approach helps to detect and overcome any ingestion issues and for the ingestion/content coordinators to be informed about the Content Providers work.

The Content Providers can simply access the helpdesk support through the reserved area of the EAGLE website (Fig. 6). WP3 and WP4 can supervise directly the activities, the issues posted in the forum and give suggestions that can be read by all the community. Furthermore, since the forum is open, also the Content Providers that have already solved an issue can suggest and share with the others the possible solutions.

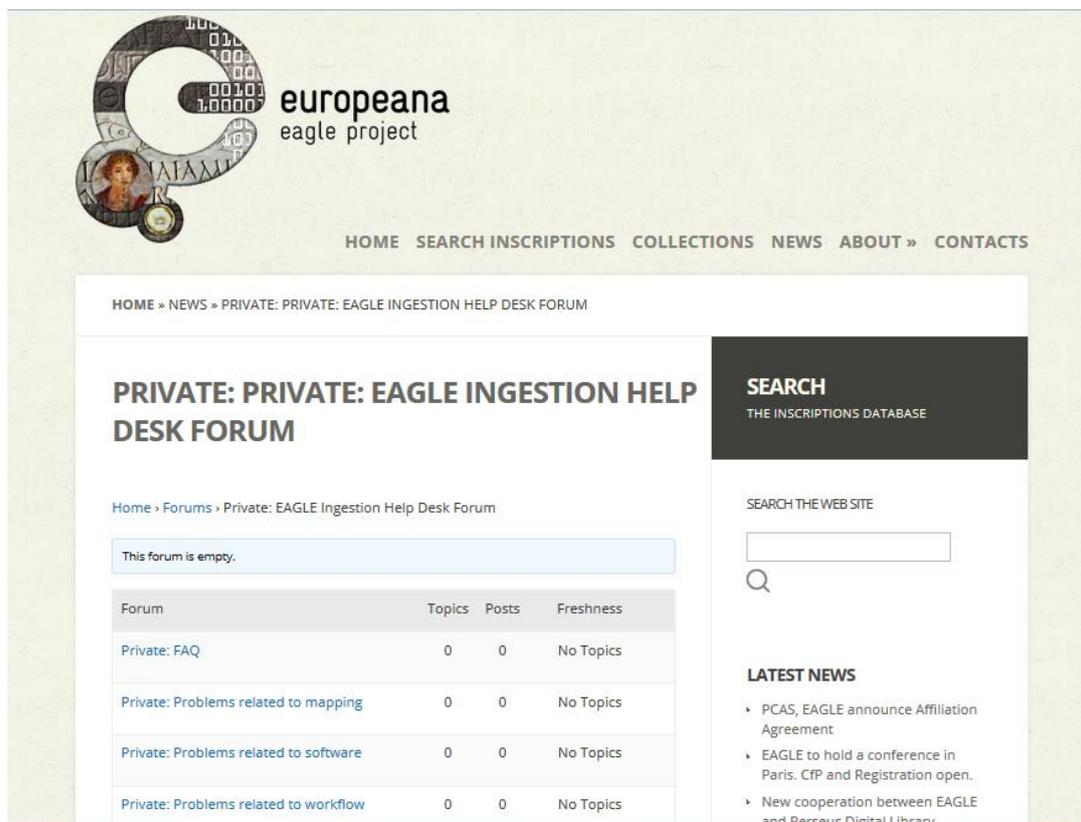


Fig. 6. The online helpdesk forum integrated in the EAGLE portal.

The forum is divided in four subject areas (Table X). A section is dedicated to problems related to mapping (Mapping): here the Content Providers can post questions or ask information about the transformation's results of the mappings from their metadata schemas into the common EAGLE metadata format and into EDM (Europeana Data Model). The second section is devoted to the problems related to the ingestion tool (Software): here the Content Providers can write in order to inform about any kind of malfunctioning, bugs, access problems, etc. The third section gathers all the problems related to the ingestion workflow (Workflow): here the Content Providers can ask clarifications or information about the ingestion steps, within EAGLE and from EAGLE to Europeana, or ask for an intermediation between them and Europeana for problems of visualization of their content in Europeana portal, and so forth. The last section is committed to collect the Frequent Asked Questions and to facilitate further the solution of common and frequent issues among the partners. All the sections are provided with a system of numbering of the topics activated in each section, of the posts and of the updates.

FORUM SUBJECT AREA	TOPICS
FAQ	The section will be dedicated to the Frequent Asked Questions
MAPPING	Information/questions about the results of the mappings in the EAGLE metadata schema and in EDM
SOFTWARE	Problems with the tool (e.g. bugs, access, etc.)
WORKFLOW	Problems related to the ingestion steps, with Europeana, etc.

Table X. The sections of the helpdesk forum.

At today, the service is ready to start, together with the effective beginning of the ingestion activities. According to the activities and the needs of the Content Providers, the service could be updated and refined in order to better support all the ingestion process along the project time span.

1.10.4 Relations with the other Work Packages

Another important matter in the monitoring of the ingestion is the relation of the WP3 with the other Work Packages that are connected with the content too. In particular, WP2, WP4 and WP5. In fact, EAGLE addresses other issues complementarily, each one corresponding to a specific work package.

Concerning WP2, Work Package 3 participates in the works of the established Groups. WP2 in fact has specific working groups of partners and external experts who explore each particular issue:

- WG1: GIS and Terminologies
- WG2: Translations and content curation
- WG3: IPR and User Engagement

For what concerns the WP4, WP3 is in tight connection with the tasks of the package since it executes the aggregation workflow and they act together as intermediate between Providers and Europeanana. Data is collected from/supplied by the content providers, the aggregated and enhanced data are returned to them for check and curation, the mapped metadata records with direct links to the digital objects are supplied to Europeanana and the external discovery services.

Finally, WP3 will monitor the work of WP5 for what concerns the content to be used for the storytelling applications, the upgrade of the EAGLE portal in relation to the content made available through the work of WP3. The contribution to the piloting activities (selection and enrichment of the inscription to be used in the mobile and storytelling applications) and to the validation phase.

All these WPs are interconnected because are based on the content and on the core aim of the project: the monitoring of their work and the relations established give a further value to the results of the WP3 itself. For this reason WP3 works also on the monitoring of the Working Groups/Work Packages, in order to check the state of art and the achievements in the related fields. To do so, WP3 participate in the other WPs' and WGs' meetings in order to be kept informed of any issues that may have an impact on the general work. Thanks to this activity, it is also possible to control any potential bottlenecks or solutions of the issues that could slow down the procedure also of the partners' content provision and vice versa.

1.11 FROM EAGLE TO EUROPEANA: THE INGESTION WORKFLOW

This part of the document wants to illustrate the workflow implemented within the EAGLE project and the relations among the Work Packages involved, the Content Providers and Europeanana team. The ingestion procedures and monitoring of the WP3 are connected with the activities of the Work Package 4 "EAGLE Aggregation and Image Management infrastructure" and particularly with the technical infrastructure developed for the content ingestion to Europeanana¹⁸.

All the elements previously described in this report, are necessary for the establishment of this workflow that starts from the work of the Content Providers, passes through the activities of WP3 and WP4 and

¹⁸ For the technical specifications, please refer to the related documentation: D4.1 "AIM Infrastructure Specification" and D4.2.1 "First Release of AIM Infrastructure".

arrives to the Europeana ingestion team action, in a cyclic way in case of adjustments and update of the content ingested (Fig. 7).

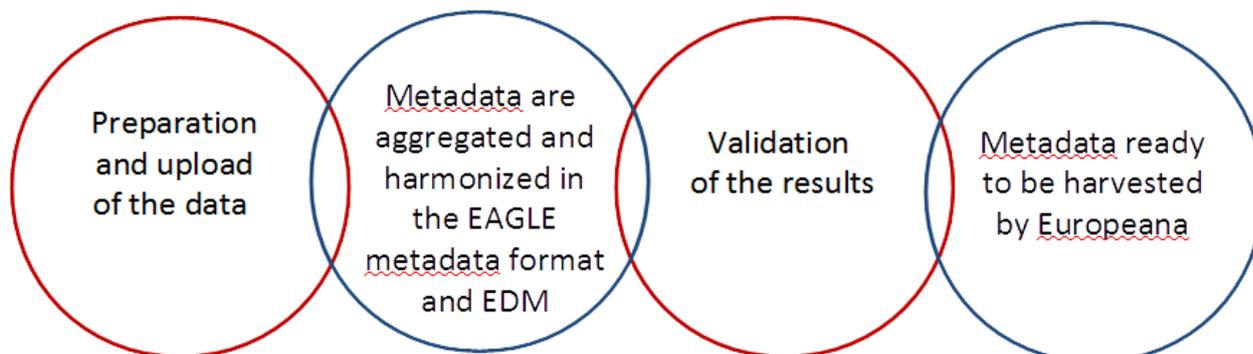


Fig. 7. The workflow from the Content Providers towards Europeana

Online availability of the resources

The first step of this procedure consists of the work of the Content Providers regarding their collections, their preparation, export and upload of the data.

Content Providers have to guarantee the online access to their collections and that they are compliant with the Europeana requirements, relatively to the Europeana Data Model specifications. In fact, Europeana can publish the Providers' data on the portal only if all the mandatory elements of the EDM are provided¹⁹. For this reason, the providers have to guarantee that these elements are supplied in their metadata schemas before the mapping into EAGLE and of course that all the resources are available online. This means that each item needs to have the URL of the metadata (that gives the possibility of linking to the landing page of the resource as represented in the Provider archive) and/or the URL to the digital object described by the metadata.

In technical terms, it is the `edm:object` (the URL of a thumbnail representing the digital object). If there isn't this thumbnail, the `edm:isShownAt` (an unambiguous URL reference to the digital object on the provider's website in its full information context) or `edm:isShownBy` (an unambiguous URL reference to the digital object on the provider's website in the best available resolution/quality²⁰). Fig. 8.

Export of the metadata in xml format

Content Providers have to export their metadata from their collection management system in XML format (Encoding: UTF-8) and send them through the export system implemented by each provider (via OAI-PMH, FTP, HTTP) or through an FTP server implemented by WP4 for those they have no export methods, to the EAGLE infrastructure. Furthermore, the majority of the providers have to guarantee that their metadata comply with EpiDoc as agreed in the Deliverable 3.1²¹. Only some Providers have not to

¹⁹ Regarding the requirements of Europeana and the mandatory fields of EDM to be provided, WP3 has elaborated a document that has been released as Annex III in the D3.1 "EAGLE metadata model specification". Please refer to the above-mentioned Deliverable for the technical specifications.

²⁰ The Europeana minimum quality requirement for the pictures is of 300 dpi.

²¹ The Providers have to guarantee the compliance to EpiDoc because it is the base of the EAGLE metadata format.

guarantee the compliance to EpiDoc since their collections will be collected in their own original metadata format and mapped to Entities/properties, identified in the EAGLE Metadata Schema²².

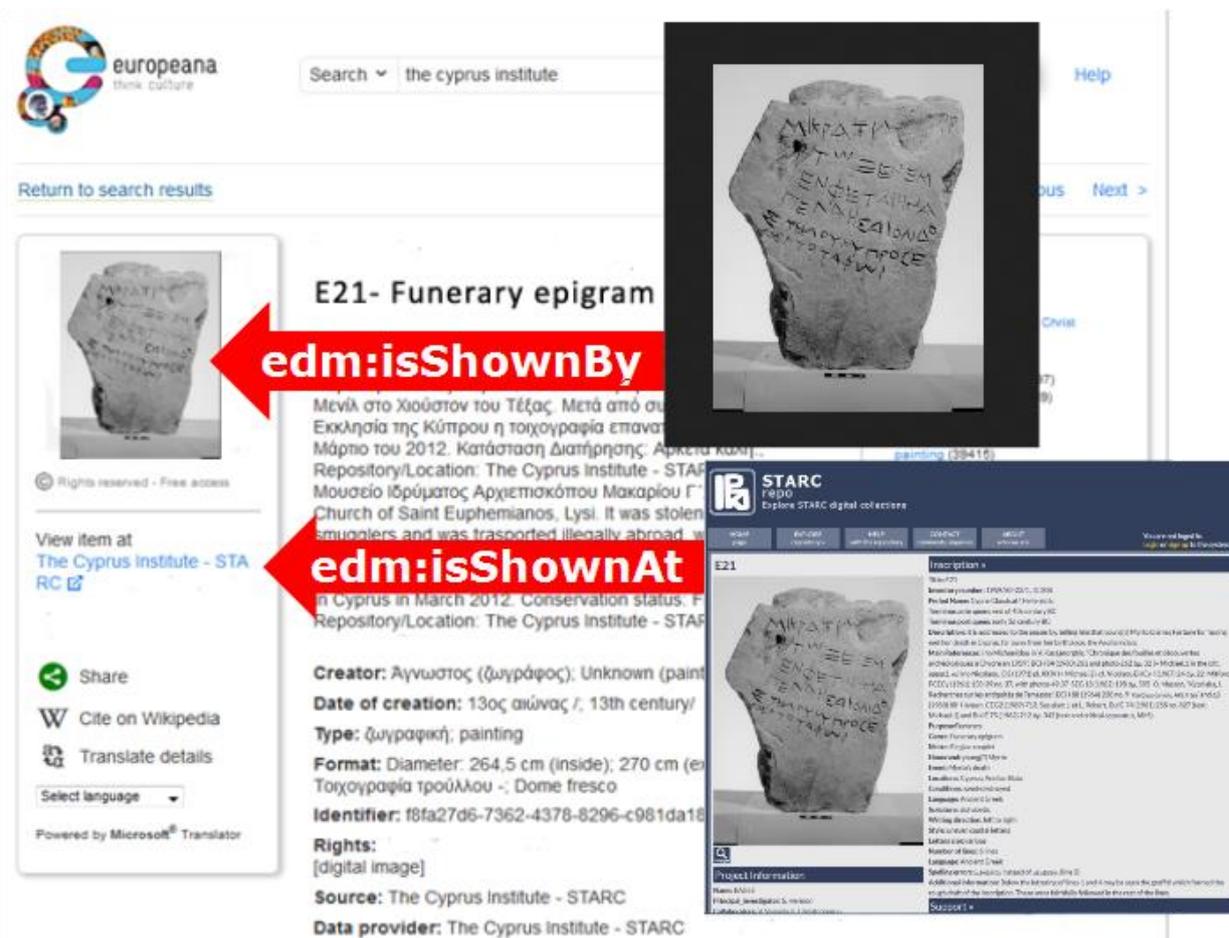


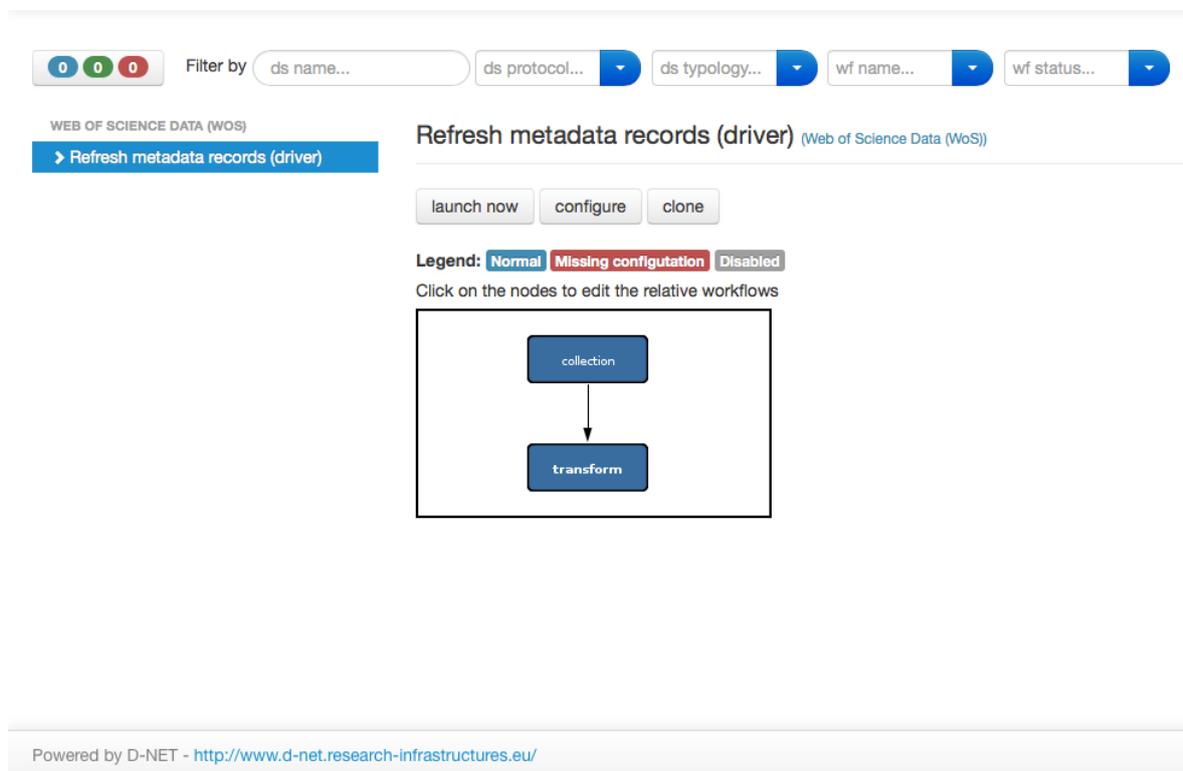
Fig. 8. Simulation of a result of the publication of a digital inscription in Europeanana with the explanation of the edm:isShownAt and edm:isShownBy.

Transformation into EAGLE metadata format and into EDM

The next steps consist of the transformation of the metadata of the Providers collections into the EAGLE metadata format. The XML metadata records provided by the Content Providers are gathered and stored into the Native Metadata Store. After that, metadata in native format are transformed into the EAGLE Metadata Format, making them structurally uniform, and stored in a Transformed Metadata Store (see Deliverable 4.2). (Fig. 9)

The mapping for the transformation into the common metadata format, within the EAGLE infrastructure, is done on the base of the specifications given by the Providers and published in the Deliverable D3.1 “EAGLE metadata model specification”. The Deliverable reports about the work done on the alignment of the different terms corresponding to each data provider to the common set of terms of the EAGLE Metadata Schema. After this passage, the EAGLE metadata schema is transformed in EDM in order to be harvested by Europeanana for the publication in its portal.

²² The metadata of these collections in fact do not fit in EpiDoc properties and they need to be mapped in EAGLE. They are: AKGD collection of the Cyl, Ubi Erat Lupa of PLUS and Arachne of DAI.



0 0 0 Filter by ds name... ds protocol... ds typology... wf name... wf status...

WEB OF SCIENCE DATA (WOS)

> Refresh metadata records (driver)

Refresh metadata records (driver) (Web of Science Data (WoS))

launch now configure clone

Legend: Normal Missing configuration Disabled

Click on the nodes to edit the relative workflows

```
graph TD; collection --> transform;
```

Powered by D-NET - <http://www.d-net.research-infrastructures.eu/>

Fig. 9. The ingestion infrastructure.

At the end of this passage performed by WP4, Content Providers have to check the results of the transformation. The EAGLE infrastructure allows the Providers to check the results through the Content Checker (Fig. 10). This tool allows to check that the files ready to be delivered to Europeana are correctly transformed and are compliant with Europeana technical requirements.

At this point, as soon as the Providers agree with the transformation, the metadata are published by WP4 in the OAI-PMH server of EAGLE and made them available for the Europeana harvesting.

Publication in Europeana

After the publication in the EAGLE infrastructure, WP3 is in charge to inform the Europeana Operations Team of the datasets that are ready for the harvesting. In agreement with the Europeana team and as already established in other Europeana sisters projects, WP3 organizes the process and the deadlines from the Content Providers side and then communicate to Europeana the datasets ready. Usually, Europeana asks that every 21st of the month the ingestion teams of the projects communicate the datasets ready to be harvested²³. For this reason, by the 21st of the month the works have to be concluded and the Content Providers have to inform WP3 about the datasets ready. Consequently, WP3 communicates to Europeana the list of the datasets. At this point Europeana works the data harvested and starts a process of quality control. After circa fifteen days Europeana sends to WP3 its validation report and WP3 informs the Content Providers and WP4 about the possible quality feedback and remedy

²³ The first EAGLE harvesting is foreseen for September 2014, but already WP3 is in contact with the Europeana Operation Team and it is agreed to test the harvesting with the first batch of data that will be published in EAGLE in April 2014.

actions to be performed on the datasets before the final publication in the portal. In this case, the procedure has to be repeated in order to adjust/fix the content. If no adjustments are needed, the dataset is finally published on the Europeana portal.

Genre :	
Person(s) :	Arthur Brauner (is created by)
Synopsis :	
Subject(s) :	
Type :	
Item(s) :	
Comment :	Text
Archive :	Deutsches Filminstitut - DIF

```
Record XML
<efg:efgEntity>
  <efg:nonAVManifestation>
    <efg:nonAVManifestation>
      <efg:identifier scheme="CP_CATEGORY_ID">DIF_1
      <efg:recordSource>
        <efg:sourceID>_node_1082497</efg:sourceID>
        <efg:provider schemeID="Institution acronym">
          </efg:provider>
        <efg:provenance>Arthur Brauner-Archiv im Deutsch
        <efg:specificType>Produktionsmaterial</efg:specificType>
        <efg:digitalFormat>application/pdf</efg:digitalFormat>
        <efg:item>
          <efg:identifier scheme="CP_CATEGORY_ID">DIF_1
          <efg:provider>Deutsches Filminstitut - DIF</efg:provider>
          <efg:country>DE</efg:country>
          <efg:aggregator>filmportal.de</efg:aggregator>
          <efg:fileFormat>application/pdf</efg:fileFormat>
          <efg:isShowable>http://www.filmportal.de/mate
          <efg:isShowableBy>http://www.filmportal.de/sites
          </efg:item>
        </efg:nonAVManifestation>
      </efg:nonAVManifestation>
    </efg:nonAVManifestation>
    <efg:identifier scheme="CP_CATEGORY_ID">DIF_Non
    <efg:recordSource>
      <efg:sourceID>_node_1082497</efg:sourceID>
      <efg:provider schemeID="Institution acronym">id=
    </efg:recordSource>
    <efg:title lang="DE">
```

```

    <efg:collection>
      <efg:collection>
        <efg:identifier scheme="CP_CATEGORY_ID">Filminstitut Austria_collection_E96803a26059e2260c8c889509904e85</efg:identifier>
        <efg:title lang="en">Archival collection</efg:title>
        </efg:collection>
      </efg:collection>
      <efg:productionYear>1917</efg:productionYear>
      <efg:title lang="en">
        <efg:geographicScope></efg:geographicScope>
        <efg:temporalScope>Erst-Hilfskämpfe in Schmalz und Eis. Aufnahmen der Filmaufnahme des k.u.k. Kriegspostquartiers</efg:temporalScope>
        <efg:relation-Original Title></efg:relation>
      </efg:title>
      <efg:keywords type="Project" lang="DA">
        <efg:term id="19171914"></efg:term>
      </efg:keywords>
      <efg:keywords type="Person" lang="DA">
        <efg:term id="Documentary"></efg:term>
      </efg:keywords>
      <efg:keywords type="Genre" lang="DA">
        <efg:term id="Kriegspostquartier"></efg:term>
      </efg:keywords>
      <efg:keywords type="Place" lang="DA">
        <efg:term id="Alpen"></efg:term>
      </efg:keywords>
      <efg:keywords type="LCSH" lang="DA">
        <efg:term id="CP_CATEGORY_ID">FAA_corporate_2e0761dd332889e69327e7615181</efg:term>
        <efg:term id="CP_CATEGORY_ID">Sechs-Film</efg:term>
        <efg:term id="CP_CATEGORY_ID">Produktion company</efg:term>
        </efg:keywords>
      <efg:keywords type="LCSH" lang="DA">
        <efg:term id="http://id.loc.gov/authorities/subjects/n2008113801">type="LCSH" lang="DA">
          <efg:term id="n2008113801">World War, 1914-1918 -- Austria</efg:term>
        </efg:term>
      </efg:keywords>
      <efg:keywords type="LCSH" lang="DA">
        <efg:term id="n79212783">Italy</efg:term>
      </efg:term>
      </efg:keywords>
      <efg:keywords type="LCSH" lang="DA">
        <efg:term id="n493007809">Mountain warfare</efg:term>
      </efg:term>
      </efg:keywords>
      <efg:keywords type="LCSH" lang="DA">
        <efg:term id="n493007809">Military maneuvers</efg:term>
      </efg:term>
      </efg:keywords>
      <efg:keywords type="LCSH" lang="DA">
        <efg:term id="http://id.loc.gov/authorities/subjects/n2008113801">type="LCSH" lang="DA">
          <efg:term id="n2008113801">World War, 1914-1918 -- Austria</efg:term>
        </efg:term>
      </efg:keywords>
    </efg:keywords>
  </efg:collection>
  </efg:efgEntity>

```

This value is invalid

Allowed values for KeywordType vocabulary are:

- Form
- Genre
- Person
- Place
- Subject
- n/a
- LCSH

Fig. 10. The Content Checker.

4. CONCLUSIONS

The current Deliverable reports about the planning and preparation of the ingestion in order to set up the workflow for the publication of the digital collections of the EAGLE Consortium in Europeana and to allow the future implementation and monitoring of the ingestion.

As conclusion, the document described the plan and organization of the content provision.

In detail, it has been assessed and updated the list and amount of the content to be delivered to Europeana and explained the changes and the actions planned by the Content Providers to cover the amount declared. The organization of the workflow and of the guidelines for the Content Providers have been reported, together with the description of the monitoring and supporting procedures set up. Moreover, the Deliverable gives a plan of the deadlines set up internally to meet the project Milestones and the next steps.

At the moment, there are no particular problems to be reported but the focus is always on the availability of the data declared and on the on-time provision to EAGLE and to Europeana. All the Providers are aware of the first ingestion Milestone (MS8) that aims at the ingestion of 375.000 items ready for the Europeana harvesting in September 2014 (month 18).

The effective content provision and ingestion towards Europeana starts now that the tool and the technical specification are ready. The ingestion platform is released in March 2014 (Deliverable 4.2) and the official ingestion starts.

The next WP3 Deliverable is due to September 2014 (month 18) in conjunction with the MS8 and it will report on the progress of the EAGLE content ingestion into Europeana, the work done by the Content Providers and of the parts in charge of the content coordination for reaching the goal.

BIBLIOGRAPHY

Piccininno M. Vassallo V., (2013), *Il flusso di lavoro nei progetti di aggregazione di contenuti culturali digitali. Buone pratiche e controllo della qualità*, in *Digitalia*, vol. 1 (2013). <http://digitalia.sbn.it/article/view/726/500>.

Vassallo V., Christophorou E., Hermon S., Niccolucci F., *Revealing cross-disciplinary information through formal knowledge representation – a proposed Metadata for ancient Cypriot inscriptions*. Conference Proceeding of Digital Heritage 2013, 28 Oct – 1 Nov 2013, Marseille.

Vassallo V., Piccininno M. (2012), *Aggregating Content for Europeana: a Workflow to Support Content Providers*, proceedings of the conference “Theory and Practice of Digital Libraries – TDPL” 2012, G. Buchanan, E. Rasmussen and F. Loizides (eds.), Springer.

ANNEX I

PARTNER ID (1)	DATABASE / COLLECTION Name	DATABASE / COLLECTION Web address	METADATA MODEL / SCHEMA / FORMAT (8)	PERSISTENT IDs yes/no	PERSISTENT IDs type	CONTROLLED VOCABULARY (10)	METADATA EXPORT PROTOCOL	LANGUAGE(S) (12)	OBJECT TYPES (13)	DESCRIPTION (14)	RIGHTS (15)
UNIROMA1	EDR - Epigraphic Database Roma	www.edr-edr.it	see EDR_metadata_schema.pdf	no	-	We have list of word that we use to fill the fields that describe the support of the inscriptions and their typology. These lists are in Latin. See http://www.edr-edr.it/English/table_en.php for details	FTP	Italian, Latin	image, text	Greek and Latin Inscriptions from Italy, Sicily and Sardinia from VII a.C. - to VII d.C.	Public domain
EDB	Epigraphic Database Bari	www.edb.uniba.it	Proprietary Schema	No	-	-	None. Implementing XML-based export.	English	.jpg	Early Christian Inscriptions of Rome (III-VII cent. CE)	
UHEI	EDH	http://edh-www.adw.uni-heidelberg.de/home	html metadata, 52 fields	not yet stable	-	- province - modern country - inscription type - inscription	FTP preferred	German, English	image, text	Greek, Latin, Punic and Iberic inscriptions from the former Roman	photographs completely publicly visible as thumbnails,

						bearer - language - material (to a lesser degree)				provinces Greek and Byzantine inscriptions of Aphrodisias	available where rights for online-publication conferred to EDH by the IPR
UOXF	Last Statues of Antiquity	http://laststatues.classics.ox.ac.uk/	A Proprietary schema was used, although it could be worked on to bring it in line with some other standard. An example is attached.	Yes	Record Identifier, of format LSA####	Yes, in the following fields: City, Province, Region, Object, Material, Re-Use, Form, Language, Text Type, Position and Gender.	None at present, can be developed further to use XML though.	English	image, text	Greek and Latin inscriptions from statue bases erected between AD 284 and the seventh century	Public Domain
UAH	HEpOI	www.eda-bea.es	proprietary Schema: relational DB with several tables	yes	internal DB record number	Glossaries for Object type, Object material. Object function, Ancient and Modern geographical names, Chronology, Repositories (i.e. Museums, collections)		Spanish, English, German (only for Admin purposes) Portuguese (in preparation)	Text data, jpg	Latin, Greek, Indigenous inscriptions	Proprietary and/or public domain
PLUS	Ubi Erat Lupa	http://www.ubi-erat-lupa.org/	Proprietary	yes	http://www.ubi-erat-lupa.org/	See database	-	German, English	Texts and	Monuments and	cc &

		erat-lupa.org	schema		.ubi-erat-lupa.org/monument.php?id=321	description			images	inscriptions	copyright
UBB	-	-	No existing metadata	-	-	-	-	Latin, Greek	jpg, tiff, pdf	Greek and Latin inscriptions from Dacia and Dobrogea	Public domain
ELTE	no	no	descriptions in text format, no metadata yet	no	-	-	-	Hungarian, German	images (jpg), text	Roman inscriptions from Brigetio	Agreements with the involved museums (Museum of Komárom, Hungarian National Museum)
UNIPU	-	-	Proprietary schema	-	-	-	-	-	-	-	-
ZRC SAZU	-	-	-	-	-	-	-	English, Slovenian	Images, maps, texts (jpg, tif, pdf)	Latin inscribed monuments from Emona and its territory	Public Domain, Rights Reserved
AUSONIUS	PETRAE	petrae.tge-adonis.fr	TEI	yes	petrae.tge-adonis.fr/id/id_insc = inscription identifier (numérotation petrae)		OAI-PMH	fr	jpg	Greek and Latin inscriptions	right reserved
DAI	iDai.Books	www.arachne.	TEI,	yes	URI	none	OAI-PMH	German, English	images,	Ancient	Creative

		uni-koeln.de	METS+MODS , RAW XML data						texts	inscriptions, sites and monuments with inscriptions	commons
DAI	inscriptions		CIDOC CRM & Dublin Core are planned	yes	URI	none	OAI-PMH	German, English			
DAI	objects		CIDOC CRM, Dublin Core, RAW XML data	yes	URI	none	OAI-PMH	German, English			
DAI	buildings		CIDOC CRM, Dublin Core, RAW XML data	yes	URI	none	OAI-PMH	German, English			
DAI	scenes		CIDOC CRM, Dublin Core, RAW XML data	yes	URI	none	OAI-PMH	German, English			
DAI	individual motifs		CIDOC CRM, Dublin Core, RAW XML data	yes	URI	none	OAI-PMH	German, English			
DAI	receptions		CIDOC CRM, Dublin Core, RAW XML data	yes	URI	none	OAI-PMH	German, English			
DAI	reproductions		CIDOC CRM, Dublin Core, RAW XML data	yes	URI	none	OAI-PMH	German, English			

CYI	STARC repository	http://public.cyi.ac.cy/starcRepo/	CyInscription/STARC metadata	no	to be implemented	no (to be implemented)	no (to be implemented)	Modern Greek, English	pdf, jpg, 3D	ancient Cypriot literature; Cypriot archaeology	Creative Commons, Public Domain
BSR	Inscriptions	www.bsrdigitalcollections.it	METS schema in xml: DC and MODS	YES	http://www.bsrdigitalcollections.it/WebArk/XMLrecords/00000/0000512.xml	TGM I & II, TGN, LOC Subject Headings, LOC Authorities for names	FTP	English	Image	Greek and Latin inscriptions from Libya (Cyrenaica) and South Etruria (Italy)	Rights Reserved
BSR	Inscriptions of Roman Tripolitania	http://irt.kcl.ac.uk/irt2009/inscr/xmlrepo.html	TEI	YES	http://irt.kcl.ac.uk/irt2009/IRT001.xml		Download from the website	English	Image	Greek, Latin and Christian inscriptions from Libya (Tripolitania)	Rights Reserved

ANNEX II

1. Provider information

Information about the organization providing content:

Name of the organisation	
Short name of the organisation	
Country	
Website	

Contact person for managerial questions:

Name	
E-mail	
Telephone	
Skype ID	

Contact person for technical questions:

Name	
E-mail	
Telephone	
Skype ID	

2. Content description

Description of the collection(s) to be provided to EAGLE. Please duplicate the table for each collection, if needed.

Collection name	<i>Please, specify the collection as in the DoW</i>
Collection URL	<i>If the collection is not online yet, please specify the reasons (e.g. digital library under re-engineering, etc.) and the estimated date of publication</i>

<p>Amount of digital objects</p>	<p><i>Amount of digital objects linked to the metadata. NB: The ratio between metadata and digital objects (1 d.o : 1 metadata; n d.o : 1 metadata)</i></p>
<p>Amount of metadata</p>	<p><i>Amount of metadata that you will aggregate within EAGLE. NB: Europeana aggregates metadata linking to digital objects</i></p>
<p>Object type(s)</p>	<p><i>Object types according to Europeana: IMAGE, TEXT, SOUND, VIDEO, 3D</i></p>
<p>Metadata export format</p>	<p><i>Please indicate whether your metadata records already comply with EpiDoc as agreed and described in Deliverable 3.1. If not, please state an estimation about when the compliance will be met. Cyprus, Lupa and images from Arachne will be collected in their own original metadata format (as they do not fit in EpiDoc properties) and mapped to Entities/properties identified in the EAGLE Aggregation Metadata Schema.</i></p>
<p>Protocol to export metadata</p>	<p><i>Please specify by which means (XML records provided via OAI-PMH, FTP, HTTP) are you intended to expose/offer your metadata to the EAGLE aggregation system. For those providers passing through EDH and EDR, please specify it. NB. More about this topic will be covered during the training session in Ljubljana, so just state your possibilities.</i></p>
<p>Amount of metadata for the first testing</p>	<p><i>In order to test the EAGLE aggregation platform, it is recommended that all the partners provide a sample of their metadata. Please specify the amount and the sample you provide for the ingestion platform testing.</i></p>
<p>Available rights</p>	<p><i>Please specify the rights that are on the digital objects. Europeana requires that data providers apply a rights statement for all digital objects described in their metadata. The rights that apply to the digital object will also apply to the previews used in the Europeana portal. The available rights statements are published at this page http://pro.europeana.eu/web/guest/available-rights-statements. Please specify if you still need to clarify your rights.</i></p>
<p>Notes</p>	<p><i>Please add any further information you think is useful.</i></p>

ANNEX III

Mid-term internal deadlines and Milestones							<i>Estimation for the MS8 on the base of the amount declared for the first ingestion and in the DoW</i>
On the base of the amount CP declared for the first internal deadline (April 2014) and in the DoW, a forecast of the provision percentage for each Provider is made until the MS8, in order to calculate the effective provision and monitor the eventual deviation in the foreseen amount.							
N.	Short Name	April 2014	May 2014	June 2014	July 2014	August 2014	September 2014
1	UNIROMA1	1.000	3.500	3.500	3.500	3.500	15.000
2	UNIBA	500	2.687	2.687	2.687	2.689	11.250
3	UHEI	1.000	3.812	3.812	3.812	3.814	16.250
4	UOXF	1	93	93	93	95	375
5	UAH	200	14.637	14.637	14.637	14.639	58.750
6	PLUS	35.805	-	-	-	-	35.805
7	UBB	800	50	50	50	50	1.000
8	ELTE	800	-	-	-	-	800
9	UNIPU	500	-	-	-	-	500
10	ZRC SAZU	108	-	-	-	-	108
11	AUSONIUS	10	310	310	310	310	1.250
12	DAI		50.000	50.000	50.000	50.000	

		50.000					250.000
13	CYI	10	25	25	25	25	110
14	BSR	1.016	-	-	-	-	1.016
	TOTAL	91.750	75.114	75.114	75.114	75.122	392.214