

DELIVERABLE

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

Report on the contributions to Europeana – second release

Deliverable 3.3.2

version: 1.0

Revision: final

Authors:

Valentina Vassallo (The Cyprus Institute)

Contributors:

Pietro Liuzzo (UHEI)

Reviewers:

Claudio Prandoni (Promoter)

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	27/03/2015	Valentina Vassallo	The Cyprus Institute	First version
0.2	27/03/2015	Pietro Liuzzo	UHEI	Contributions to the first version
0.3	09/04/2015	Valentina Vassallo	The Cyprus Institute	Integration of the reviewers' comments
1.0	10/04/2015	Claudio Prandoni	PROMOTER	Final check

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This deliverable contains original unpublished work except where clearly indicated otherwise.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 ROLE OF THE DELIVERABLE 3.3.2 IN THE PROJECT	7
2. CONTENT PROVIDERS AND COLLECTIONS.....	8
2.1 CONTENT PROVIDERS	8
2.1.1 New affiliated partners.....	9
2.1.2 Data Exchange Agreement (DEA) and EAGLE agreements.....	10
3. THE INGESTION PLAN	13
3.1 THE INGESTION PLAN FORM	13
3.1.1 Content assessment.....	14
4. IMPLEMENTING AND MONITORING THE INGESTION.....	15
4.1 THE INGESTION WORKFLOW	15
4.2 MONITORING OF THE MID-TERM INTERNAL MILESTONES	17
4.3. FEEDBACK.....	17
4.3.1 Direct contacts	18
4.3.2 Technical meetings.....	18
4.3.3 The EAGLE File Transfer Protocol	19
4.3.4 The D-Net metadata record inspector (EAGLE infrastructure)	20
4.3.5 Europeana feedback	21
4.4. METADATA ENRICHMENT	22
4.4.1 Duplicates identification	23
4.4.2 Vocabularies.....	23
4.4.3 Translations	24
5. REPORT ON THE CONTRIBUTIONS TO EUROPEANA	25
5.1. INSCRIPTIONS AND VISUAL REPRESENTATION: THE EUROPEANA PUBLICATION.....	27
5.2. QUALITY TASK-FORCE	29
6. CONCLUSIONS	30

LIST OF ABBREVIATIONS

ABBREVIATION	EXPLANATION
CP	Content Provider
DoW	Description of Work
EDM	Europeana Data Model
WP	Work Package
WG	Working Group
DEA	Data Exchange Agreement
TM number	Trismegistos number
CHO	Cultural Heritage Object

EXECUTIVE SUMMARY

This Deliverable, created by the Cyprus Institute (Cyl), with inputs of WP2 (UHEI) cited and explained in the text, reports about the second release of the contribution to Europeana. Particularly focuses on the progress of the ingestion of EAGLE content towards Europeana, providing updates in the workflow and in the current ingestion.

The document is structured in 6 Chapters.

Chapter 1 is an introduction about the objectives of this Deliverable and its role in the project. It is an overview of the objectives and tasks of WP3, focussing also on its Milestones.

Chapter 2 is dedicated to the Content Providers and their collections, with a specific focus on the new affiliated partners. Among them, the new contributors of data to the EAGLE project and to Europeana are highlighted. Moreover, it is given the status of the Data Exchange Agreement (DEA) signed by the Providers for the data publication into the European portal, as well as the status of the EAGLE agreements signed by the new affiliated partners for the provision of content (e.g. translations, pictures for Wikimedia) to EAGLE project.

Chapter 3 is dedicated to the ingestion plan and the importance of implementing and monitoring the data ingestion. Particularly, information about the content assessment and the data amount that will be sent to Europeana, through EAGLE project, is provided.

Chapter 4 is the summary of the workflow developed to monitor the ingestion, the different sources of information used to perform it and the collaboration with others WPs for the completion of the Tasks.

Finally, Chapter 5 is dedicated to the report on the contributions to Europeana, with a detailed description of the work done by the Content Providers at today, the new data ready to be harvested by Europeana and the first results of the publication in the portal. In the last part, a discussion about a possible further increase of the EAGLE data quality is provided.

Chapter 6 draws the conclusions and gives information about the next steps of the Work Package 3.

1. INTRODUCTION

This report gives an overview of the activities carried out by Work Package 3 during the period 1st October 2014 – 31st March 2015. As already expressed in the first release of this Deliverable (D 3.2.1), the work reported here is strictly connected with the Deliverable 3.2, since they are parts of the same work aimed at the development of a methodology and of a workflow for the EAGLE content contribution into Europeana. The report shows all the steps of the workflow: the Content Providers' planning of the content is followed by the progress of the ingestion, the feedback collected and the monitoring of the contributions activities towards Europeana. All the topics already treated in the previous issue (Deliverable 3.3.1 'Report on the contributions to Europeana – first release') will be cited but not repeated in this Deliverable unless they are useful for the general comprehension or there are important changes and updates in the structure.

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 is 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 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 ('Definition of the EAGLE metadata model' and 'Metadata mapping'); K.U.LEUVEN ('Duplicates identification') and UNIROMA1 ('Metadata enrichment and contextualization'). All the Content Providers are involved in the Work Package contributing 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 progres and to assure that the Tasks and Milestones will be achieved.

The Milestones connected to this Work Package 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 3.3.2. 'Report on the contributions to Europeana – second release' illustrates the ongoing aggregation process, focussing on the monitoring of the ingestion and the data contribution to EAGLE and to Europeana. The cooperation between WP2 and WP4 is fundamental to carry out the WP3 tasks.

This document reports about the work carried out from month 19 (October 2014), to month 24 (March 2015).

1.2 ROLE OF THE DELIVERABLE 3.3.2 IN THE PROJECT

The aim of the Deliverable 3.3.2 is the report on the contribution from the EAGLE Content Providers to Europeana. This Deliverable is important to assess the implementation and monitoring of the ingestion workflow. As the previous version, from now onward all the Deliverables will report on the effective contribution of the content to the EAGLE infrastructure and to Europeana. They are strictly connected with the achievement of the ingestion procedure and of the Milestones: 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 reports the schedule of the WP3 Deliverables and Milestones (in light blue, the current report period):

MONTH	12	18	24	30	36
DELIVERABLE	D3.2	D3.3.1	D3.3.2	D3.3.3	D3.3.4
MILESTONE	-	MS8	MS12	MS15	MS16

Table I. WP3 Deliverables and Milestones.

2. CONTENT PROVIDERS AND COLLECTIONS

2.1 CONTENT PROVIDERS

As described in the Description of Work, the project Consortium 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 users.

The list of the EAGLE participating partners and Content Providers are reported in Table II.

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.

The EAGLE content is constituted of inscriptions (texts) connected to digital images (photographs, drawings, scanned notes, manuscripts and archive material, ancient books) and their related metadata. At present, there are no changes or updates regarding the type of collections that will be provided to the project with respect to what has been reported in D3.2.

2.1.1 New affiliated partners

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 are encouraged to join the Best Practice Network and to contribute their content. All the content providing partners fulfill an important role in awareness-raising and recruiting new content providers across Europe.

WP3 assists WP2 to monitor 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 monitors the involvement of the stakeholders that will contribute with content to EAGLE and to Europeana¹.

With respect to what we reported in September 2014, regarding the interest of various institutions and initiatives to collaborate with EAGLE and to provide content to the project, at today 12 affiliated partners signed the agreement with the EAGLE project².

1. Pontificia Commissione di Archeologia Sacra (contributes photos to EDB which in turn aggregates the content to Europeana)
2. The Israel Museum, Jerusalem (contributes items related to the EAGLE project)
3. Attic Inscriptions Online (contributes translations to the EAGLE MediaWiki)
4. University of Beograd (contributes photos, uploading them to Wikimedia Commons and linking them to the partner databases)
5. Città di Anagni (contributes photos)
6. King's College London (contributes the Inscriptions of Aphrodisias collection)

¹ The methodology and the activities about the enlargement of the network are described in the Deliverable D.2.1 'Networking infrastructure and terms of reference'.

² They correspond to 14 collections. In the list there are cases in which the content has already been provided by the institutions even if the agreement documents have not yet arrived to the project coordinator.



7. Centre for the Study of Ancient Documentation (UOXF, new projects not initially included in the project) including : Roman Inscriptions of Britain (RIB), Monumenta Asiae Minoris Antiquae XI (MAMA XI), Vindolanda Texts
8. Ashmolean Museum Latin Inscriptions project (UOXF)
9. Historical Institute of Montenegro (contributes photos, uploading them to Wikimedia Commons and linking them to the partner databases)
10. Iscrizioni Latine Archaiche (ILA)
11. Inscriptions from Sicily
12. Inscriptions of Greek Cyrenaica

Among the new affiliated partners, some have expressed their interest to contribute to the project and have already started to put their data at disposal of the project. The data of these institutions are under processing in order to comply with the EAGLE and Europeana specifications. They will be included in the next Europeana's harvestings.

Other institutions declared their interest to participate to the project: for these organizations the affiliations' agreements are currently under discussion and/or preparation. The possible interested institutions are to be finally confirmed and at the moment they are:

- Inscriptiones Graecae of the Louvre
- Inscriptiones Graecae
- US Epigraphy project
- Topoi project Berlin Humbolt University
- EPNNet
- CLE Galliae – Hispaniae
- Inscriptiones from Israel-Palaestinae
- Inscriptions from Attica, Dalmatia and Pompei from the SFB993 project
- Ilia State University
- Epigraphic Database Clauss Slaby

2.1.2 Data Exchange Agreement (DEA) and EAGLE agreements

Within the WP3 tasks, an important requirement for the content provision and its monitoring is the sign of the Europeana Data Exchange Agreement by all the EAGLE Content Providers.

As explained in the Deliverable 3.2 "Ingestion plan" and in the successive D.3.3.1 'Report on the contributions to Europeana – first release', every Content Provider supplying content to Europeana has to subscribe the Agreement. Therefore the WP3 task (together with WP1) 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. Since September 2014 all the EAGLE Content Providers signed the DEA and their contents are therefore allowed to be transmitted to Europeana for the publication in the portal.

As reported in the paragraph 2.1.1, new affiliated partners joined the project. In this phase of the project WP3 is following the signature of the DEA by the new Content Providers (curated by WP2) and the current situation is as follows in Table III:

Organisation	Providing content to Europeana	DEA	Providing content to EAGLE	EAGLE agreement
Pontificia Commissione di Archeologia Sacra	Yes (via EDB)	Signed		
The Israel Museum	Yes (content related to the EAGLE material)	Signed		
Attic Inscriptions Online			Translations (only in the EAGLE Media Wiki)	Signed
University of Beograd			Photos to Wikimedia Commons (photos are not yet online and will be uploaded directly to Wikimedia Commons)	Signed
Cittá di Anagni	Yes (10 photos)	Signed		
Roman Inscriptions of Britain	Yes (2.401 items)	Signed (under University of Oxford)		
Inscriptions of Aphrodisias (KLC)	Yes (1.505 items)	Not signed yet		
Centre for the Study of Ancient Documentation	Yes (386 items)	Signed (under University of Oxford)		
Inscriptiones Graecae	Yes (text and translations, undeclared amount)	Waiting for approval		
US Epigraphy project	Yes	Waiting for approval		

Iscrizioni Latine Archaiche (ILA)	Yes (80 items)	Signed		
Topoi project Berlin Humbolt University	Yes (texts of christian inscriptions, undeclared n°)	Waiting for approval		
AshLi project	Yes (300 inscriptions)	Signed		
Historical Institute of Montenegro			Photos to Wikimedia Commons (photos are not yet online and they will be uploaded directly to Wikimedia Commons)	Signed
Organisations that declared interest to participate but they have still to be confirmed				
EPNet	Yes (ca. 37.000 items)		(to be confirmed)	
CLE Galliae – Hispaniae	Yes (amount not declared)		(to be confirmed)	
Inscriptiones from Israel-Palaestinae	Yes (ca. 1730 items)		(to be confirmed)	
Inscriptiones from Attica, Dalmatia and Pompei from the SFB993 project	Yes (ca. 1000 items)		(to be confirmed)	
Ilija State University	Yes (1 item)		(to be confirmed)	

Table III. The new affiliated partners and the sign of the DEA for the content provision to Europeana and EAGLE agreements for the content provision to the project.

3. THE INGESTION PLAN

The tasks of the WP3 started at M1 (April 2013) and will last for the whole duration of the project (M36, March 2016)

The ingestion plan consists of all the activities on which is based the EAGLE aggregation procedure and the content delivery to Europeana: surveys, technical sessions, partners support, procedures, methodologies, feedback and milestones. WP3 aims at the planning and preparing the ingestion and at its implementation and monitoring.

The task T3.3. “Planning and preparing the ingestion”, concluded in March 2014, is connected with T3.4. “Implementing and monitoring the ingestion”, where 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.

The task of the WP3 leader is to ensure that the Content Providers make their content available according to the project schedule and that they provide the content declared. Therefore, WP3 leader, from month 12 (March 2014) started to operate the monitoring of the content ingestion into Europeana and its progress. From this Deliverable onwards, every 6 months WP3 reports about this monitoring activity and on the contribution to Europeana.

3.1 THE INGESTION PLAN FORM

As described in the Deliverable 3.2 “Ingestion plan”, in order to perform the WP3 Tasks of keeping under control the content delivery to Europeana and to report about contributions as declared in the Description of Work (DoW), a working tool has been elaborated. The ingestion plan form is in fact useful to keep trace of the Content Providers information about the items they will provide to Europeana, along all the project period.

The tool is used to confirm or update the content declared in the EAGLE Description of Work, to describe the status and the availability of the collections to be ingested and to trace the ingestion progress. The ingestion plan form is maintained updated according to the new deadlines and Milestones. In the Table IV the updated document is provided with the new fields related to the periods to be monitored.

No	Short Name	DEA	DoW	Second Survey	MS8 (September 2014)	MS12 (March 2015)
1	UNIROMA1	✓	60.000 images (photographs and drawings)	60.000
2
...
TOTAL			1.472.198

Table IV. The ingestion plan form

The ingestion plan form data are cross-checked together with the information gathered with other working tools and coming from the feedback of all the project's partners. The result is a synthesis of all the analytical reports of the Content Providers' contributions to EAGLE and Europeana.

3.1.1 Content assessment

The ingestion could be influenced by technical and administrative changes within the partner institutions. As experienced in other Europeana sisters projects, sometimes the digital content committed in the DoW is no more available (for different reasons) and needs to be replaced. In other cases some providers may be late in content provision. Since these issues could have an impact on the final content provision, it is important to keep under control the total provision.

The first update and control of the ingestion plan is the result of two Surveys launched and performed in March 2013 and January 2014. Considering the importance to keep under control the general amount of the EAGLE contribution to Europeana, in January 2015 WP3 launched another Survey (more appropriately a *tour de table*) for the Content Providers, in order to assess the total amount of each Provider and understand if the project is in line with the contribution foreseen in the DoW, or if an amendment is needed. The *Tour de table* is included in the activities of the Task 3.4 'Implementing and monitoring the ingestion' for the effective ingestion procedure scope and the update of the information declared by the Content Providers in the Description of Work, as well as to monitor the content to be provided to Europeana. The aim of this activity 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 are committed to give to the project. The Providers have been asked to give information about the collections amount 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.

Compared to the previous information collected in September 2014, in January 2015 no changes were reported by the Content Providers in the total amount of their collections.

In March 2015, and in particular in the occasion of the Plenary meeting in Cyprus (10th March 2015), last minutes differences have been declared by few Content Providers. Particularly:

- UAH declared to have some difficulties in the collection, preparation and enrichment of their data that are causing a delay in their ingestion procedure (difficulties that are due to copyright issues and to the complexity of their metadata structure that needs more time and work to be mapped in the EAGLE and Europeana metadata models).
- The exact amount to be provided by AUSONIUS is still under definition.

The remaining Content Providers confirmed or increased the amount committed in the DoW and their contribution is in line with the agreed time schedule.

It has to be noted that the other Content Providers that during the 2014 Survey declared less content (UNIBA, PLUS, BSR), at today they confirmed that they solved their problems and that they can contribute the overall amount as declared in the DoW.

4. IMPLEMENTING AND MONITORING THE INGESTION

This section is an update of the same section included in D3.3.1. It is fully reported again here in order to provide an overall picture of the process.

In Task 3.4 'Implementing and monitoring the ingestion', each Content Provider applies the best practices and implements locally the necessary workflows and procedures to ensure the continuous provision of metadata, surrogates and digital content for ingestion to the EAGLE Aggregator (WP4) and to Europeana.

The WP3 leader ensures the monitoring of the progress of the ingestion of new content into Europeana as planned in T3.3. The aim of this task is:

- to ensure that the work of aggregation keeps pace with the releases of Europeana;
- to ensure that the work of metadata mapping keeps pace with the releases of Europeana;
- to ensure that the work of ingestion keeps pace with the releases of Europeana.

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 maintained unchanged as declared in the Description of Work or as updated by the providers during the assessment (surveys, direct contact, emails, etc.).

4.1 THE INGESTION WORKFLOW

This section is a summary of the workflow implemented within the EAGLE project, and it explains the relations between the Work Packages involved (particularly WP3 and WP4), the Content Providers and Europeana ingestion team³. The workflow starts from the work of the Content Providers, passes through the activities of WP3 and WP4 and arrives to the Europeana ingestion team action, in a cyclic way in case of adjustments and update of the content ingested (Fig. 1).

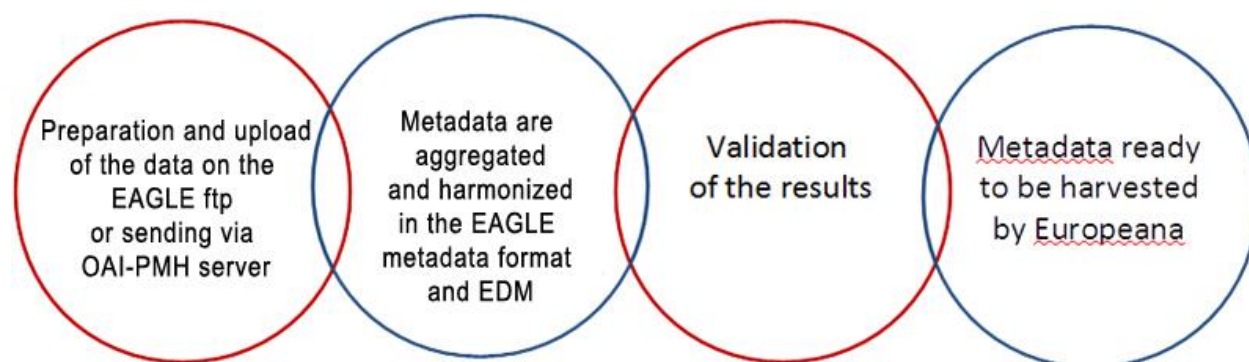


Fig. 1. The ingestion workflow

The workflow consists of four steps and is useful to follow how the aggregation and monitoring process is being carried out within EAGLE and which cooperation mechanisms among the project WPs, Content Providers and Europeana is being set up.

³ For the full description, see the Deliverable 3.2 'Ingestion plan'.

- First step: preparation, export and upload of the data.

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 EAGLE and Europeana requirements. 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. Content Providers have to export their metadata from their collection management system in XML format and send them through the export system implemented by each provider, if any (via OAI-PMH, FTP, HTTP), or through the FTP server implemented by WP4 for those they have no export methods, to the EAGLE infrastructure.

- Second step: data aggregation and harmonization

The second step is represented by the aggregation platform developed by CNR-ISTI, leader of the WP4. The platform is a customized version for EAGLE of the D-Net software. It is the web environment that allows WP4 to upload the project Content Providers data, map to EAGLE metadata schema and to EDM in order to send the data to Europeana.

The ingestion service enables:

- the data uploading of the Content Providers' data;
- the mapping and transformations of the metadata records into EAGLE records and the aggregation in the repository;
- the validation of the content;
- the mapping and transformation of the EAGLE data into EDM and the final transmission to the Europeana ingestion via OAI-PMH.

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 CPs⁴. After this step, the EAGLE metadata schema is transformed in EDM in order to be harvested by Europeana⁴ for the publication in its portal.

- Third step: validation of the results

Content Providers have to check the results of the transformation⁵. The Task 3.2 of preparing the necessary mappings for the metadata conversions in order to allow the alignment of proprietary provider's metadata schemata with the EAGLE schema, started in January 2014 and will last until March 2016. This means that Content Providers have to keep under control the results of their mapping and provide a continuous validation of the results and update of them in case of necessity. WP3, responsible for the metadata mapping (the University of Alcalà is in charge of leading this task) together with WP4 and the Content Providers, will maintain updated the metadata mappings until the end of the project. WP3, WP4 and the Content Providers, carry out the work of control, validation and revision of the mappings and data for the publication in the EAGLE infrastructure. The feedback procedure helps to achieve the results.

During the last Plenary meeting (10 March 2015, Nicosia) a proposal of establishing a quality task force for the data published both in the EAGLE infrastructure and in Europeana, has been discussed (see Paragraph 5.2 of this Deliverable).

⁴ The specifications are published in the Deliverable D3.1 "EAGLE metadata model specification".

⁵ The EAGLE infrastructure allows the Providers to check the results and allows to check that the files ready to be delivered to Europeana are correctly transformed and are compliant with Europeana technical requirements.

- Fourth step: EAGLE publication and Europeana harvesting

The last step consists of the mapping and transformation of the EAGLE metadata schema into EDM and the consequent Europeana harvesting. After the transformation in fact the metadata are published by WP4 in the OAI-PMH server of EAGLE and made them available for the Europeana harvesting.

4.2 MONITORING OF THE MID-TERM INTERNAL MILESTONES

According to the ingestion workflow, WP3 asked to all the Content Providers to contribute to the March deadline (Milestone 12), distributing their work during the period October 2014-March 2015⁶.

It has been asked to the Content Providers to declare in advance the amount of data that they were going to ingest until March 2015, in order to monitor their work during this period. The internal intermediate deadlines are aimed in fact at a continuous monitoring and assessment, to check the work done and plan the next steps. The comparison of the results achieved with the objectives and milestones foreseen is of great importance for the evaluation of possible bottlenecks, to prevent them and to look for the right solution on time.

In this phase of the ingestion it has to be reported that only very few Content Providers were able to plan a monthly ingestion. In the majority of the cases the Providers are not able to foresee in advance the amount of material available by each month and they prefer to declare a total amount by the deadline of the Milestone. This doesn't mean that the Content Providers are not working during the all period but simply they prefer to set a total final amount that then they split according to the availability of the content during the time-span. In fact, the monitoring of the provision, carried out thanks to several feedback typologies and further detailed in the following paragraphs, demonstrates that the Content Providers work and provide the content, in some cases during the period and in other cases all at the end of the period.

In Chapter 5 'Report on the contribution to Europeana' it is reported the effective provision of all the Content Providers to the EAGLE infrastructure and ready to be harvested by Europeana for the final publication.

4.3. FEEDBACK

The collection of feedback is an important step of the workflow in the contributions to Europeana and in the monitoring of the ingestion procedure. It is in fact part of the monitoring and supporting activities of the Providers' work and of all the actors involved in the contribution to Europeana. This feedback activity consists of different sources of information in order to:

- collect information from the users to face their problems in the aggregation process;
- answer content providers' requests to overcome their difficulties;
- define an effective ingestion methodology and its monitoring.

The feedback collection methodology is part of the general methodology developed for the contribution of data to Europeana. Feedback is important because it helps WP3, WP4 and the Content Providers

⁶ 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.

themselves to refine and improve the ingestion process and to support the contribution of this medium/large amount of content to Europeana.

The feedback is gathered making use of several tools and sources of information, particularly:

- direct contacts with the Content Providers
- training sessions and face-to-face meetings
- the EAGLE FTP
- the Metadata Record Inspector of the EAGLE infrastructure
- Europeana feedback

Thanks to these source of information, the progress of the ingestion is documented and recorded in the working tools elaborated by WP3 (and shared with WP4), where it is possible to trace the collected information. The 'Report on the contribution to Europeana' makes use of all these elements and of the crosscheck of the information from the different tools and actors participating in the project.

4.3.1 Direct contacts

As reported in the Deliverable 3.2 'Ingestion plan', WP3 implemented direct relations with the Content Providers. 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. Together with the direct communications, the EAGLE mailing list are used to inform about the general requirements, deadlines and get information about Content Providers general feedback.

All the results of these communications are reported in the working tool elaborated by WP3 to follow the Content Providers' ingestion from the beginning of the project (uploading of the content in the EAGLE infrastructure) until the end (the publication in Europeana) in order to monitor the available content and the progress of the project. The working tool is the document that helps to maintain updated the information and the communication between WP3 and Content Providers, drawing a personal history in a way fast to read and to consult⁷.

4.3.2 Technical meetings

In the occasion of the EAGLE Plenary meetings (24-25 October 2013, Pisa; 18 February 2014, Ljubljana) WP3 and WP4 organised dedicated sessions to train the Content Providers on the EAGLE metadata schema, on the use of the ingestion platform and on the ingestion procedure and aggregation to Europeana. The aim of these meetings is to keep informed the Content Providers about the technical progress of the project and to give them direct assistance about the technical specifications needed for the delivery of the content to Europeana.

Also in occasion of the last Plenary meeting in Nicosia, Cyprus, (10 March 2015), WP3, WP4 and the Content Providers had the possibility to check the status of the ingestion and discuss about the technical requirements, changes and adjustment in the light of the fulfilment of the second Milestone (MS12)⁸.

⁷ For a detailed description of the monitoring file, please refer to Deliverable 3.2 'Ingestion plan'.

⁸ The next technical meeting, if no further meetings will be needed in the meantime, is planned to be held in Bari, during the next EAGLE plenary meeting.

As always, all the material produced for the Content Providers and all the technical presentations are updated and shared in the EAGLE project repository.

4.3.3 The EAGLE File Transfer Protocol

A dedicated EAGLE File Transfer Protocol (FTP) was established to upload data from the Content Providers that do not have at disposal an OAI, FTP or web APIs to enable WP4 to collect data in a programmatic way. Therefore, pushing content to the EAGLE FTP server substituted the practice to send the content through e-mail or packages, and allowed to have a real time update of the content and of possible lack of data during the substitution of the content in the EAGLE infrastructure operated by WP4.

After a test with a couple of Providers, a guide to upload content to the EAGLE FTP server has been prepared by CNR-ISTI and shared in the EAGLE project repository together with all the needed instructions⁹. The Guide is divided in two sections: the first one is dedicated to the prerequisites needed to upload content; the second one is a step-by-step guide, enriched with instructions and images, to support the Providers during the installation of the tool and the upload and update of the data.

The rest of the Content providers send content via their OAI-PMH servers or FTP protocols.

Through the access to the EAGLE FTP server, WP3 (as well as WP4 and the Content Providers themselves) is able to monitor the upload of the content carried out by each Content Provider (Fig. 2). Every Content Provider can upload its content in its folder; WP3 can visualize the amount of files sent and the last modification done. In this way, WP3 has a clear overview of the amount of data that the Providers prepared and keep trace of their work¹⁰.

⁹ The FTP Guide is available at the following address of the EAGLE Repository: <http://goo.gl/LajFSy>.

¹⁰ The control of the data amount of the Providers that send their content through OAI server is carried out monitoring them directly.

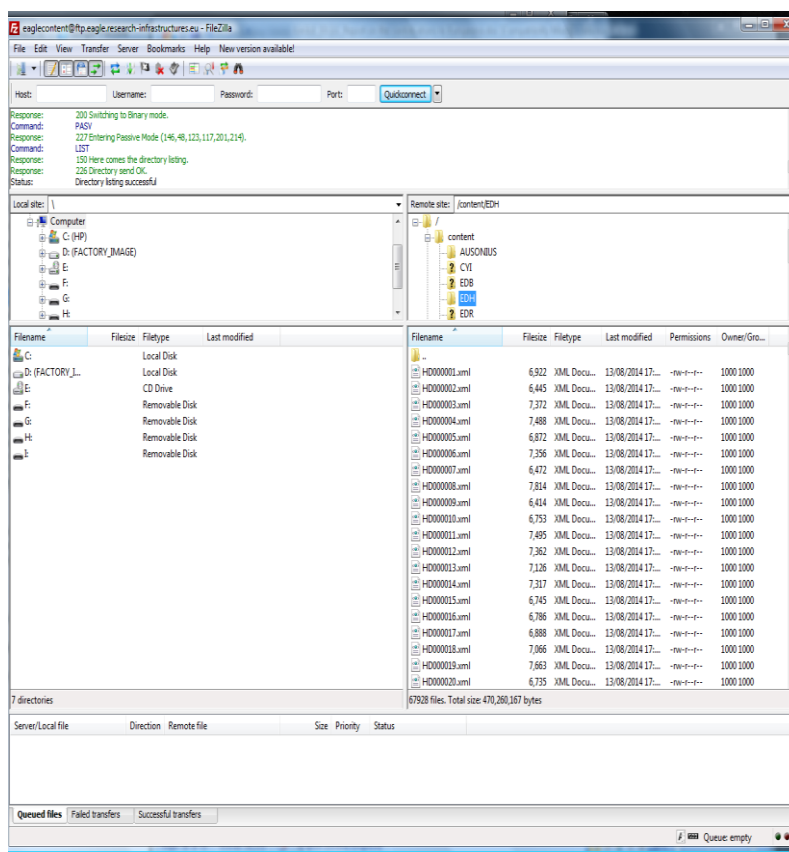


Fig. 2. The EAGLE FTP server

4.3.4 The D-Net metadata record inspector (EAGLE infrastructure)

This source of information is very useful to get an analytical overview of the effective content ingestion.

The D-Net data report functionalities provides features that allow to make detailed reports both for the single content providers and for WP3 and WP4 coordinators: the partners (after accessing with credentials) can have a quick view on the amount of metadata uploaded and transformed in the EAGLE infrastructure; the coordinators can supervise the work of all contributors. The data report offers to the users the breakdown of the items uploaded in the tool provider by provider and the consequent transformation operated after the mapping to the EAGLE metadata schema (Fig. 3).

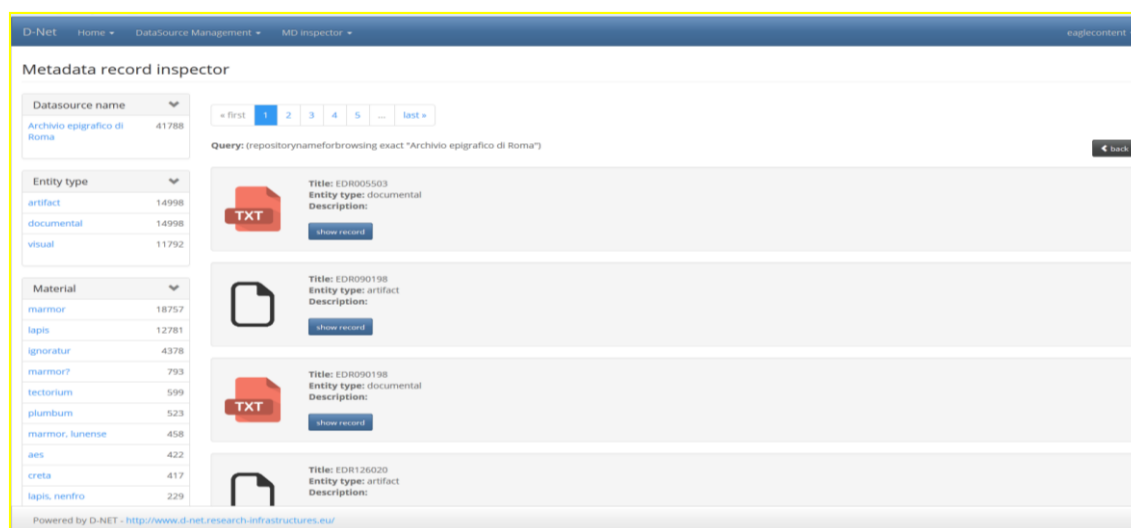


Fig. 3. The metadata record inspector: the amount of data of a Content Provider

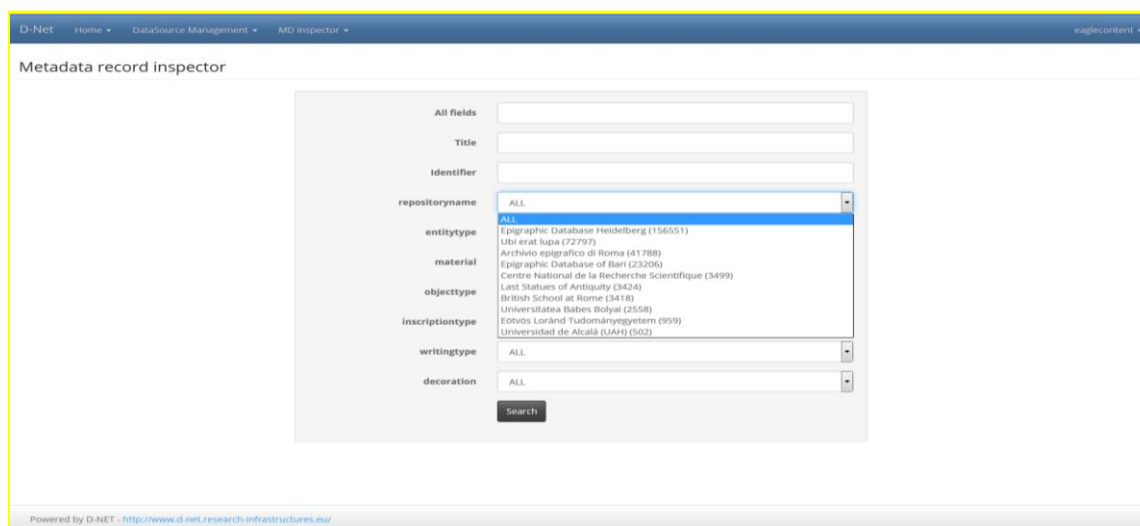
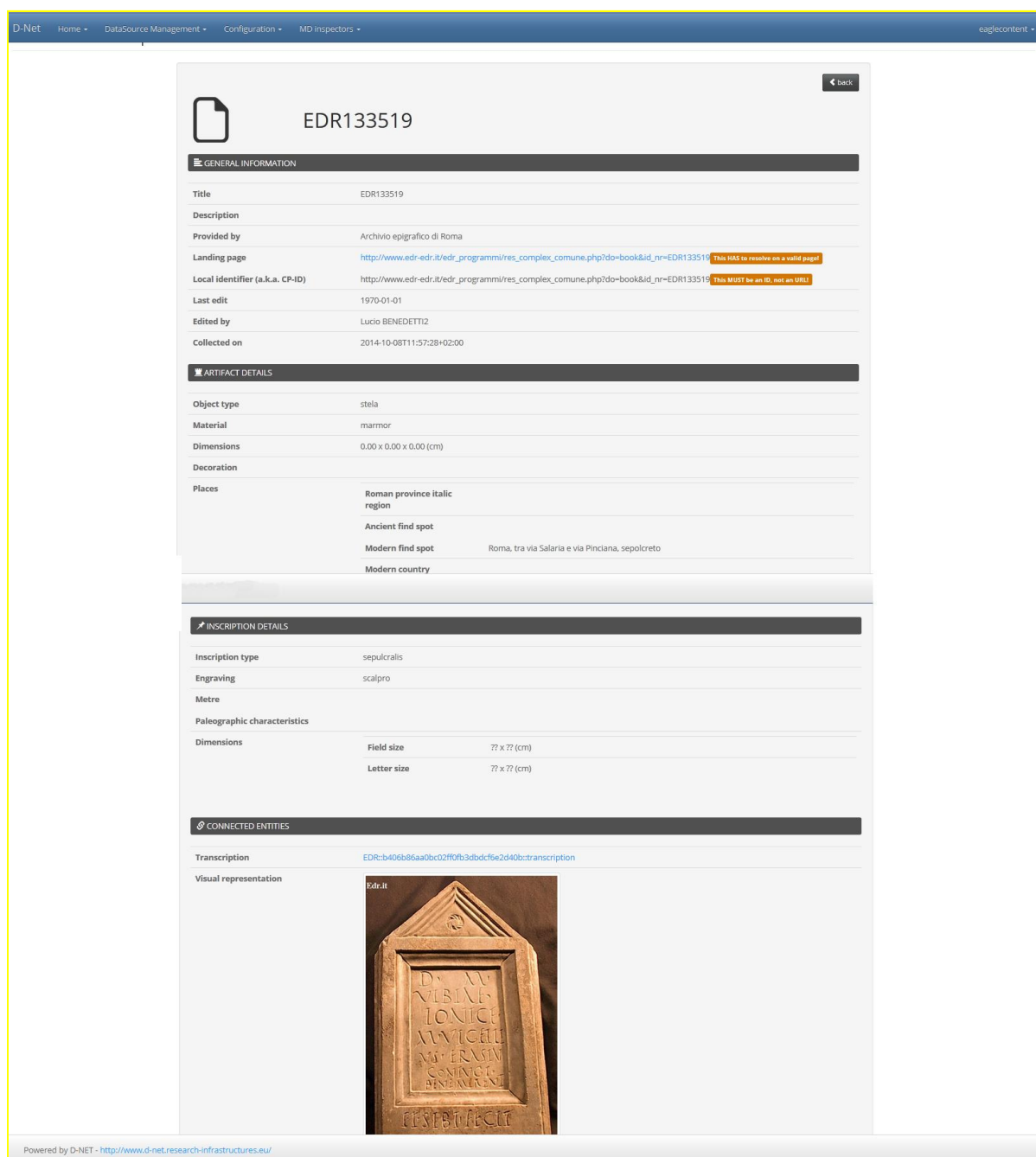


Fig. 4. The metadata record inspector in D-Net: the list of the institutions that aggregate data for EAGLE Project.

Finally, the last functionality recently added to the metadata record inspector is the visualization of the metadata details, with the generation and visualization of all the element wrappers that constitute the metadata record (Fig. 5). In this way the CP can operate a quality control of the data and easily identify, through a user-friendly interface, the content provided and eventually adjust or enhance the data.

This feature include also the possibility to check the correlations between the various items (artefacts, documental manifestations and visual manifestations, allowing the Content Providers to verify that the data for the EAGLE portal and for Europeana are correct.



The screenshot shows a web interface for a digital record. At the top, there is a navigation bar with 'D-Net', 'Home', 'DataSource Management', 'Configuration', and 'MD inspectors'. The main content area is titled 'EDR133519' and contains several sections:

- GENERAL INFORMATION:**
 - Title: EDR133519
 - Description:
 - Provided by: Archivio epigrafico di Roma
 - Landing page: http://www.edr-edr.it/edr_programmi/res_complex_comune.php?do=book&id_nr=EDR133519 (with a warning: 'This HAS to resolve on a valid page!')
 - Local identifier (a.k.a. CP-ID): http://www.edr-edr.it/edr_programmi/res_complex_comune.php?do=book&id_nr=EDR133519 (with a warning: 'This MUST be an ID, not an URL!')
 - Last edit: 1970-01-01
 - Edited by: Lucio BENEDETTI2
 - Collected on: 2014-10-08T11:57:28+02:00
- ARTIFACT DETAILS:**
 - Object type: stela
 - Material: marmor
 - Dimensions: 0.00 x 0.00 x 0.00 (cm)
 - Decoration:
 - Places: Roman province italic region, Ancient find spot, Modern find spot (Roma, tra via Salaria e via Pinciana, sepolcreto), Modern country
- INSCRIPTION DETAILS:**
 - Inscription type: sepulcralis
 - Engraving: scalpro
 - Metre:
 - Paleographic characteristics:
 - Dimensions: Field size (?? x ?? (cm)), Letter size (?? x ?? (cm))
- CONNECTED ENTITIES:**
 - Transcription: EDR:b406b86aa0bc02f0fb3dbdcf6e2d40bc:transcription
 - Visual representation: A photograph of a stone inscription with Latin text, labeled 'Edr.it'.

At the bottom, it says 'Powered by D-NET - <http://www.d-net.research-infrastructures.eu/>'.

Fig. 5. The visualization of the metadata details and content

4.3.5 Europeana feedback

The feedback from Europeana is the last important step to ensure that the content provided to the portal complies with Europeana's requirements and that is ready to be published.

The technical communications between the Europeana ingestion team, WP3 and WP4 and the feedback received by Europeana about the mapping fine-tune from EAGLE to EDM, allowed to start the final step of the ingestion procedure and arrive to the data harvesting by Europeana side.

After the publication in the EAGLE infrastructure, in fact 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 organises the process and the



deadlines from the Content Providers side and then communicate to Europeana the datasets ready to be harvested.

Another important aspect of the feedback procedure is the validation report that Europeana sends once the datasets of the providers have been harvested. In fact, Europeana harvests the datasets published in the project OAI-PMH server 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 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.

This feedback is also important to calculate the effective amount of data published in the Europeana portal after the ingestion workflow.

In accordance with the Europeana ingestion team, the harvesting is every 21st of the month. The WP3 leader, on the base of other projects' experience, decided to set up an internal anticipated deadline for the uploading of the EAGLE items on the EAGLE ftp server (or via Providers' ftp and OAI-PMH servers), in order to be on time for the effective data harvesting by Europeana and to check possible issues during the data processing in D-Net. In this way, the Consortium is ready in providing their data and WP3 to send the list of datasets every 21st of the month. In the next 15 days Europeana works the data and, if there are no further feedback for fixing issues, the datasets harvested will be published online in the portal (Table V). During the period under report, many tests of harvesting were made as well as exchange of feedback and communication between Europeana, WP3 and WP4, in order to refine the procedure and the quality of the data to be published¹¹.

Upload on the EAGLE FTP server (internal deadline)	WP3 sends to Europeana te datasets ready to be harvested	Europeana publishes the records in the portal
15th of the month	21st of the month	ca. 15 days after the harvesting

Table V. Deadlines for the Europeana harvesting

4.4. METADATA ENRICHMENT

As reported in the DoW, the EAGLE repository will be populated by metadata records following the EAGLE schema through the aggregation service. All the content-related activities aim to check and edit the aggregated metadata records through the web based editor developed in WP4, thus preparing them for external applications (e.g. the two Flagship Applications developed in WP5) to get and process them. Content providers will be able to browse their contributed items, filter them based on specific elements and values and group edit the resulted item set in order to perform:

- data cleansing (correct typographical errors, conform with typographical conventions etc.);
- data reconciliation (align elements with authority files, vocabularies or thesauri, link resources to external data sources and semantic binding);

¹¹ It has to be reported that during December 2014 Europeana did not operate any harvestings.

- manual enrichment (link and explore external data sources to extract and reuse information about resources, i.e. locations, time-spans, agents, associated with provider's items;
- contextualisation (link the epigraphs to monuments, sites, ancient topography and any available related information in order to see them in their broader historical and cultural context and to reconstruct the history of collections and the transfer of inscriptions from one place to another).

Particularly, the metadata enrichment and contextualization is operated through the following steps:

- a check of the content through the D-Net metadata record inspector (see Section 4.3.4)¹²;
- duplicates identification
- enrichment through the use of the EAGLE controlled vocabularies developed in WP2
- enrichment through the translations developed in WP2.

4.4.1 Duplicates identification

One of the methodology to enrich the metadata that will be provided to the EAGLE project is the duplicates identification through the Trismegistos platform (for a detailed description, please refer to Deliverable 3.3.1).

The task (T3.5) started at month 1 (April 2013) and officially concluded at month 12 (March 2014). At present, most of the work has been completed and only some minor refinements are still ongoing (few records to be disambiguated; final implementation of doubles and one to many/many to one correspondances in the datasets). It has to be reminded that out of the total records, only the text on the basis of literature could be disambiguated and therefore the total numbers are different from the total amount of items provided, where the Content Providers give both texts and artifacts without text (not subjected to disambiguation)

This Deliverable reports the same situation as the previous version (see Deliverable 3.3.1).

4.4.2 Vocabularies

Metadata enrichment is also performed through the use of the controlled vocabularies that have been developed by the WG1 in WP2. Some information and further specifications/clarifications are added in this Deliverable respect to the previous one.

The Working Group supported a best practice methodology which further stakeholder could use for the alignment of medium sized heterogeneous terminologies.

The EAGLE Vocabularies are the following and are maintained to always contain the maximum information available from the current content providers:

- Type of Inscription: <http://www.eagle-network.eu/voc/typeins.html>
- Object Type: <http://www.eagle-network.eu/voc/objtyp.html>

¹² The EAGLE infrastructure, beyond being the aggregator of the providers content and the WP3 working tool for the ingestion monitoring, through the metadata record inspector allows the CPs to check the correctness of their data and to further edit it.



- Material: <http://www.eagle-network.eu/voc/material.html>
- Execution: <http://www.eagle-network.eu/voc/writing.html>
- Decoration: <http://www.eagle-network.eu/voc/decor.html>
- State of Preservation: <http://www.eagle-network.eu/voc/statepreserv.html>
- Dating Criteria: <http://www.eagle-network.eu/voc/dates.html>

The alignment of the vocabularies was carried out using simple Google spreadsheets following on a previous international effort by the EpiDoc Community agreed by WG1 members. This allowed to keep always track of relevant modification (e.g. new entry and related IDs), and to open up in an easy and familiar way to any interested parties the contribution to the alignment. The Vocabularies are now fully integrated in the EAGLE portal (<http://www.eagle-network.eu/resources/vocabularies/>) and are accessible in Linked Open Data for any users¹³. Usability functions have been also added in order to facilitate use and further searches.

This stylesheets allow the matching of contents in the skos vocabulary with the string content of the data provided, so that the dereferenced terms bring the user to a content page where further connections can be explored. For an example of the skos sample, please refer to the Deliverable 3.3.1.

For new content providers a process of extraction and enlargement of the vocabularies is carried out by UHEI so that all new terms are aligned and new ones are aligned according to the guidelines in Deliverable 2.2.1.

4.4.3 Translations

A further methodology to perform the metadata enrichment is the use of the inscriptions' translations produced within the project (Task 2.2 in WP2). Translations in the Mediawiki are entered from sources outside the consortium and from new partners. The Mediawiki (available at <http://www.eagle-network.eu/wiki>), also allows to enter data:

- from registered independent users
- from users submitting translations via Perseids integration, a editorial vetting system developed in collaboration with Perseus Project
- from batch uploads with pywikipediabot (<https://github.com/EAGLE-BPN/eagle-wiki>)

All data are available to any users, to the EAGLE portal and to Content providers via a RESTful API native of Wikibase, the extension used for this purpose (<http://www.eagle-network.eu/wiki/api.php>)

To facilitate the alignment to the data in the EAGLE aggregator each translation has at least one identifier of the content providers and wherever possible a TM identifier.

¹³ The EAGLE BPN followed the same approach that the Pleiades and Pelagios project choose for geographical data and adapt it to seven other possible ways of linking data for the international community, available according to Linked Open Data good practices.

5. REPORT ON THE CONTRIBUTIONS TO EUROPEANA

The aim of the WP3 is coordinating the content provision of the partners to Europeana through the EAGLE infrastructure. For this reason, WP3 operates a continuous monitoring about the data to be provided, checking of any possible delays or problems that may have a negative impact on the content provision (see Chapter 4 for a detailed description of the procedure).

During the period under report, all the Content Providers worked to meet the deadline and sent the content to the EAGLE Aggregator in order to upload and map/transform their data in the technical infrastructure. They prepared and adjusted their data in order to export the content and make it compliant with EAGLE and Europeana specifications. Few Providers are still facing some technical issues and they are working to solve them. The majority of the Content Providers sent the planned amount (50% of the total amount as described in the DoW) through FTP or via OAI-PMH servers to EAGLE. Some of them uploaded even more items than what was planned. In some cases the Providers already met the final deadline of the project, overcoming the target of the Milestone 12 (50% of the total amount).

The content provided by the Consortium, due to the transformation into the EAGLE metadata model, will produce three different groups of metadata according to the digital objects they are connected to: artefacts, documental manifestation, visual representation. Artefacts (i.e. the information related to the physical carrier of the inscription, including the text of the inscription) and visual manifestations (i.e. the images) are those who are sent to Europeana (see Paragraph 5.1 for further details). So far in the EAGLE OAI-PMH server, already harvested or ready to be harvested by Europeana, there are 344.890 artifacts and 676.721 visual representations, for a total of 1.021.611 items.

The Table VI presents the status of the ingestion at March 2015, comparing what each Content Provider actually delivered with what it was planned for this second Milestone (MS12), i.e. 50% of the total amount foreseen in the DoW. All the issues encountered by the Content Providers are reported in the last column. Among them, the most important ones are:

- The contribution of partner UAH, who is facing several problems related to copyright issues and to the complexity of their metadata schema who is difficult to be mapped to EAGLE and Europeana models.
- The contribution of partner Cyl, whose mapping to the EAGLE and Europeana models is presently under finalisation (however, their items are ready).

All the other issues are under control.

No.	Short Name	Planned amount (MS12)	Actual amount	Details	Comments
1	UNIROMA1	30.000	94.268	61.117 artefacts and 33.151 visual representations	They are temporarily covering also the amount of UNIPU and ZRC-SAZU, while their technical issues will be solved (see below)
2	UNIBA	22.500	36.604	32.082 artefacts and 4.522 visual representations	

3	UHEI	32.500	83.545	67.998 artefacts and 15.547 visual representations	
4	UOXF	750	7.127	2.836 artefacts and 4.291 visual representations	
5	UAH	117.500	66.303	26.134 artefacts and 40.169 visual representations	Delay due to copyright issues related to the content of about twenty file-cards cabinets in which are stored data from inscriptions found in Portugal and Spain and to the complexity of the organization of the file cards archive
6	PLUS	25.500	63.621	24.016 artefacts and 39.605 visual representations	
7	UBB	2.000	1.864	1.293 artefacts and 571 visual representations	Few items are missing but they will be soon added
8	ELTE	750	825	440 artefacts and 385 visual representations	
9	UNIPU	1000	0	0	The items have been sent to EDR, but for technical problems their ingestion is postponed. Meanwhile they are replaced by 1000 items from EDR
10	ZRC SAZU	200	0	0	The items have been sent to EDR, but for technical problems their ingestion is postponed. Meanwhile they are replaced by 200 items from EDR
11	AUSONIUS	2.500	2.682	1.332 artefacts and 1.350 visual representations	
12	DAI	500.000	662.361	126.620 artefacts and 535.741 visual representations	
13	CYI	221	0	0	135 items have been uploaded but the transformation into EAGLE metadata is still to be finalised
14	BSR	678	2.411	1.022 artefacts and	

				1.389 visual representations	
	TOTAL	736.099	1.021.611	344.890 artefacts and 676.721 visual representations	MS12 achieved

Table VI. Report on the contribution to Europeana for the MS12.

During the present reporting period, two batches of content has been communicated and harvested by Europeana:

- The first batch has already been published in the Europeana portal.
- The second batch is still under publication due to Europeana’s internal delays and for issues to be fixed in the EAGLE datasets.

Before the final publication in the Europeana portal, the datasets, after have been processed, are temporarily kept online in the ingestion portal, the online working lab of Europeana (Fig. 6). The Europeana ingestion team provided WP3 and WP4 leaders with the access to this ingestion portal, in order to visualize the data before the effective visualization.

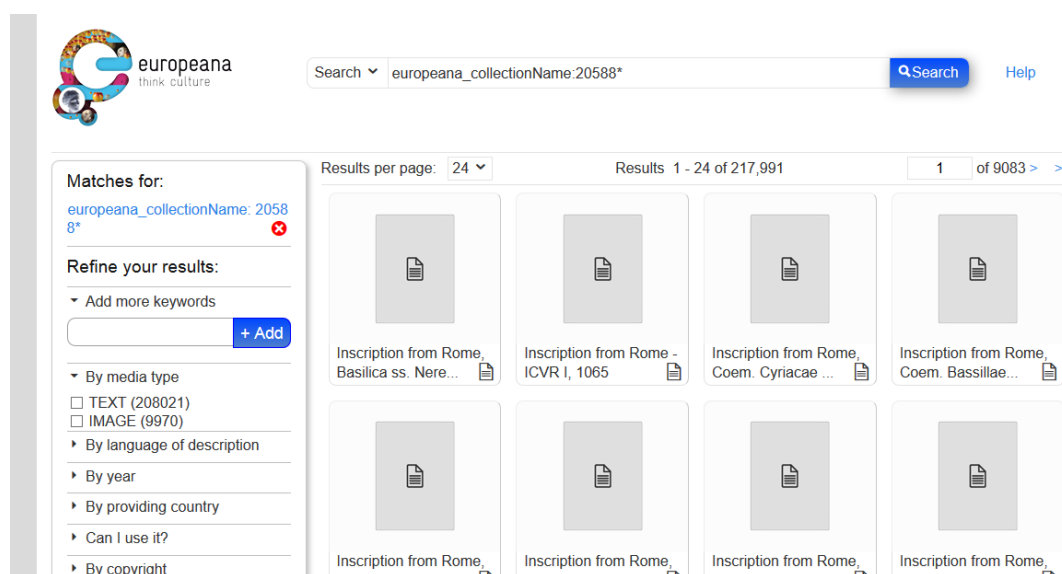


Fig. 6. The EAGLE data published in the Europeana ingestion portal

5.1 Inscriptions and visual representation: the Europeana publication

As explained in the previous section, the content provided by the EAGLE archives, due to the transformation into the EAGLE metadata model, is split in different groups of metadata according to the digital objects they are connected to: artefacts, documental manifestations and visual representations. According to what Europeana accepts to be published in its portal, the data that are sent to Europeana belong to the category of the artefacts (the inscriptions represented as texts, called by Europeana “CHO”

or “Cultural Heritage Object”) and of the visual representation (the images related to the inscriptions, called by Europeana “WebResource”). The distinction between CHO and WebResource has been introduced as a result of the introduction of the new EDM schema to provide the users with a better navigation and data retrieval and to aggregate under the same umbrella all the resources available for a Cultural Heritage Object, avoiding data duplication in the portal. For this reason, the WebResources are not anymore explicitly counted in Europeana in the numbering of the items, as they are linked to the CHO through the ore:Aggregation field, however they have their own meaning as they bring along their metadata and rights. Therefore, in order to calculate the total amount of items published in the portal, the Europeana ingestion team provides WP3 and WP4 leaders with the result of the statistics concerning the number of the CHOs and of the WebResources for each Content Provider. Table VII presents the total figures calculated by Europeana for each EAGLE Content Provider in February 2015.

Data Provider (Europeana IDs)	Number Provided	Number WebResource	Total¹⁴
Epigraphic Database Roma - EDR (europeana_collectionName:2058806*)	14.999	11.792	26791
Epigraphic Dabatase Bari - EDB (europeana_collectionName:2058807*)	22.956	1.552	24508
Epigraphische Datenbank Heidelberg - EDH (europeana_collectionName:2058801*)	67.994	22.580	90574
Last Statues of Antiquity - UOXF (europeana_collectionName:2058803*)	1.717	0	1717
Hispania Epigraphica Online - HEpOI -UAH (europeana_collectionName:2058808*)	25.435	0	25435
Ubi erat lupa - PLUS (europeana_collectionName:2058805*)	23.814	39.333	63147
Universitatea Babes Bolyai - UBB (europeana_collectionName:2058810*)	0	0	0
Eötvös Loránd Tudományegyetem - ELTE (europeana_collectionName:2058809*)	440	0	440
PETRAE - AUSONIUS - AUSONIUS (europeana_collectionName:2058804*)	1.258	1.008	2266

¹⁴ In the Europeana ingestion portal at February 2015.



British School at Rome - BSR (europeana_collectionName:2058802*)	1.022	1.390	2412
Total	159.635	77.655	237290

Tab VII. Europeana statistics: amount of CHOs and WebResources published in the Europeana portal

As it can be seen from the figures of the Table VII, there is a difference between the Europeana statistics and the number of the items published by each Content Providers in the Table VI. This is due to the fact that the statistics refer only to the first publication in the Europeana portal.

5.2 Quality task force

The content provided by EAGLE matches the current strategy of Europeana of moving from quantity to quality. This is because the national institutions participating in the project, thanks to their authoritative and comprehensive knowledge of the provided materials, are in the best conditions to make EAGLE a trusted source for cultural heritage. The EAGLE Consortium is therefore working hard to reach high quality standards that will be visible in the final results.

In general, the quality of the data that will appear in Europeana is managed by the single Content Providers that decide what kind of content to ingest and publish. The WP2, WP3 and WP4 leaders are instead in charge of the control from a different point of view that involves: the use of the vocabulary, the completeness of the metadata and the technical requirements.

Following the discussions had during the last plenary meeting in Cyprus (10 March 2015), the Consortium proposed to set up a mechanism for the definition of quality check procedure aimed to identify who is responsible for what is published in each phase of the aggregation workflow, in order to supply Europeana with good and quality checked data. The establishment of a task force to define guidelines for checking quality of the aggregated content is therefore under discussion. Of course, WP3 and WP4 leaders, as well as Europeana, has little influence over the content made available through the partner institutions. However, the role of the aggregators is to provide best practices and recommendations to assure the best quality content and to monitor this process. A possible scenario could be to set up the task force composed both by epigraphers and by technical persons: in this way it will be possible to check the data from a content and technical point of view. This task force can add another quality control layer to the one already operated by each Content Providers, cross checking the information with an external and more critical point of view.

5. CONCLUSIONS

This Deliverable is the result of the collaborative work done by WP2, WP3, WP4 and the Content Providers for the implementation of the data ingestion and it gives information about the status of the EAGLE contribution to Europeana.

The document summarises the ingestion workflow and the activities and sources of information elaborated and used to keep under control the ingestion itself. Particularly, the tools and the methodology used to quantify the contributions of the single Content Providers and to perform the final assessment of the content to be delivered to Europeana.

During the present reporting period all the Content Providers worked hard to reach the project's milestones and 1.021.611 items are now available in the EAGLE OAI-PMH server already harvested or ready to be harvested by Europeana: some are proceeding smoothly, others already completed the total provision (even if some adjustments are needed). A few Content Providers are facing some difficulties in reaching the planned target, while others gave even more content respect to the expected one. WP3 keeps monitoring the content ingestion closely.

New affiliated partners signed a Cooperation Agreement with EAGLE and some institutions declared their interest to contribute with new content.

The periodic report about the contributions to Europeana will be updated every 6 months: the next issue is planned at the end of September 2015, in conjunction with Milestone 15.