ATHENA
Access to cultural heritage networks across Europe

**ECP-2007-DILI-517005**

**ATHENA**

# Implementation plan for taking the content into Europeana

| | |
|---|---|
| **Deliverable number** | *D5.5* |
| **Dissemination level** | *Public* |
| **Delivery date** | *30 April 2011* |
| **Status** | *Final* |
| **Author(s)** | *Marzia Piccininno (MiBAC)*<br>*Maria Teresa Natale (MiBAC)*<br>*Valentina Vassallo (The Cyprus Institute)* |

***e*Content*plus***

This project is funded under the eContentplus programme[1],
a multiannual Community programme to make digital content in Europe more accessible, usable and exploitable.

[1] OJ L 79, 24.3.2005, p. 1.

# Table of Contents

**OBJECTIVE OF THE DELIVERABLE** ..... **3**

**1. INTRODUCTION: 'HOUSTON, WE HAVE LIFT OFF!'** ..... **4**

**2. WORKFLOW STEP-BY-STEP** ..... **6**

STEP 1 – CREATION OF THE ATHENA NATIONAL CONTACT POINTS 6
STEP 2 – INVESTIGATION OF THE DIGITAL COLLECTIONS 7
STEP 3 – INGESTION PLAN 7
STEP 4 - THE ATHENA INGESTER 9
STEP 5 – TRAINING AND TRAINING MATERIAL 15
STEP 6 – ATHENA HELP DESK 16

**3. CRITICISMS** ..... **17**

COLLECTION CRITICISMS 18
THE EUROPEANA AGREEMENTS 18

**4. LESSON LEARNT** ..... **20**

**5. MASTERPIECES** ..... **21**

BELGIUM 21
CZECH REPUBLIC 22
GERMANY 23
ISRAEL 31
LITHUANIA 32
RUSSIA 33

**ANNEX I: INGESTION PLAN** ..... **34**

## Objective of the deliverable

As stated in the ATHENA Description of Work, the final step of the work to be carried on by WP5 is the definition of 'the ways of taking the content into Europeana' and the drawing up of an implementation plan to be defined in cooperation with Europeana.

Since the first project meetings, it became evident that the work to be outlined in deliverable 5.5 should have been anticipated, otherwise the objective of WP5, keeping under control the content contributed by partners for the implementation in Europeana, would have been unfeasible.

Deliverable 5.5 summarises the whole work carried out until the end of the project in cooperation with other WPs, WP7 in particular as far as the technical features for aggregating the content are concerned. For this reason, some information may look redundant since it was already provided in other deliverables produced by WP5 and other WPs; repetitions will be avoided as much as possible using cross-references (highlighted in grey colour).

By the way, WP7 has to supply at month 30 a deliverable with a similar name (D7.5 'Implementation plan and access to content of museums through Europeana'); the leaders of both WPs agreed the respective tables of content in order to avoid overlapping. Deliverable 5.5 is more focused on workflow and partner contribution management, deliverable 7.5 on how the ingestion happens by the technical point of view.

# 1. Introduction: 'Houston, we have lift off!'

"Houston, we have lift off!" This was the enthusiastic expression of satisfaction of Susan Hazan of the Israel Museum of Jerusalem, when the content provided to ATHENA finally appeared in Europeana after months of work.
It was November 2010 when the first set of 1,800,000data from Germany, Italy and Israel were published in the European portal and the Mask of Gu (FIG. 1) became the ATHENA lucky charm.

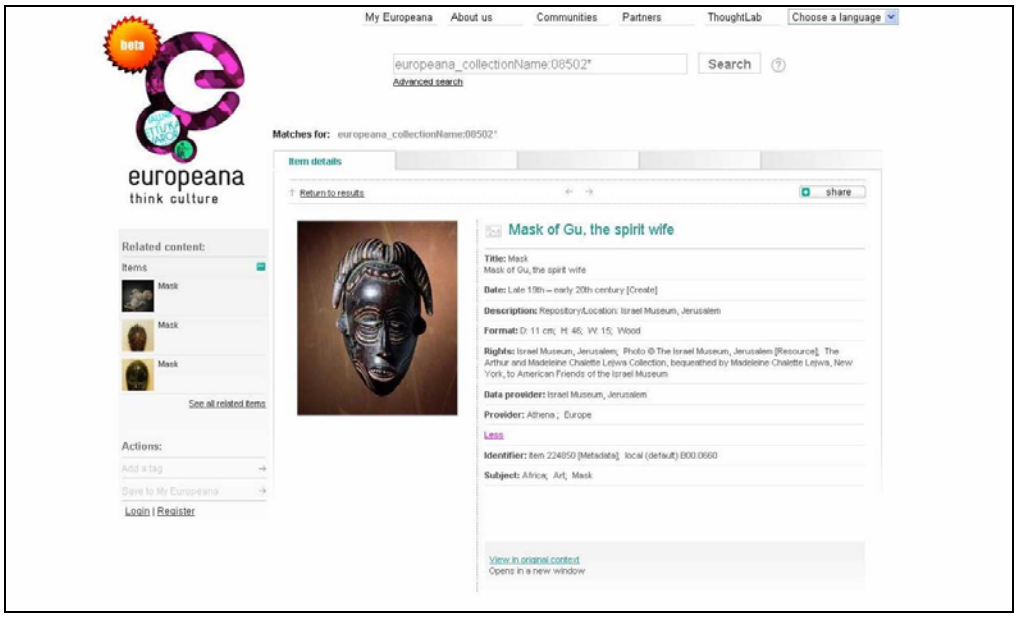



FIG. 1 - Mask of Gu, the spirit wife (Late 19th – early 20th century).
The Israel Museum, Jerusalem. This item was in the first set of data harvested by Europeana.

The enthusiasm for the success of this first stage did not come only from those who were personally involved with the publication of their metadata in Europeana. Numerous testimonials and congratulations arrived from other content providers, eager to see their online content soon: "What a thrill to see him live!", "Big news with the publication, congratulations and thanks for the hard work. The institutions that have participated will be the judges of the quality of results".

These results have been increased over time. Currently over 3.7 M data provided by the ATHENA consortium are in Europeana and other 700,000 about are waiting for the harvesting and publication (FIG. 2).

**ATHENA is the largest Europeana aggregator.**

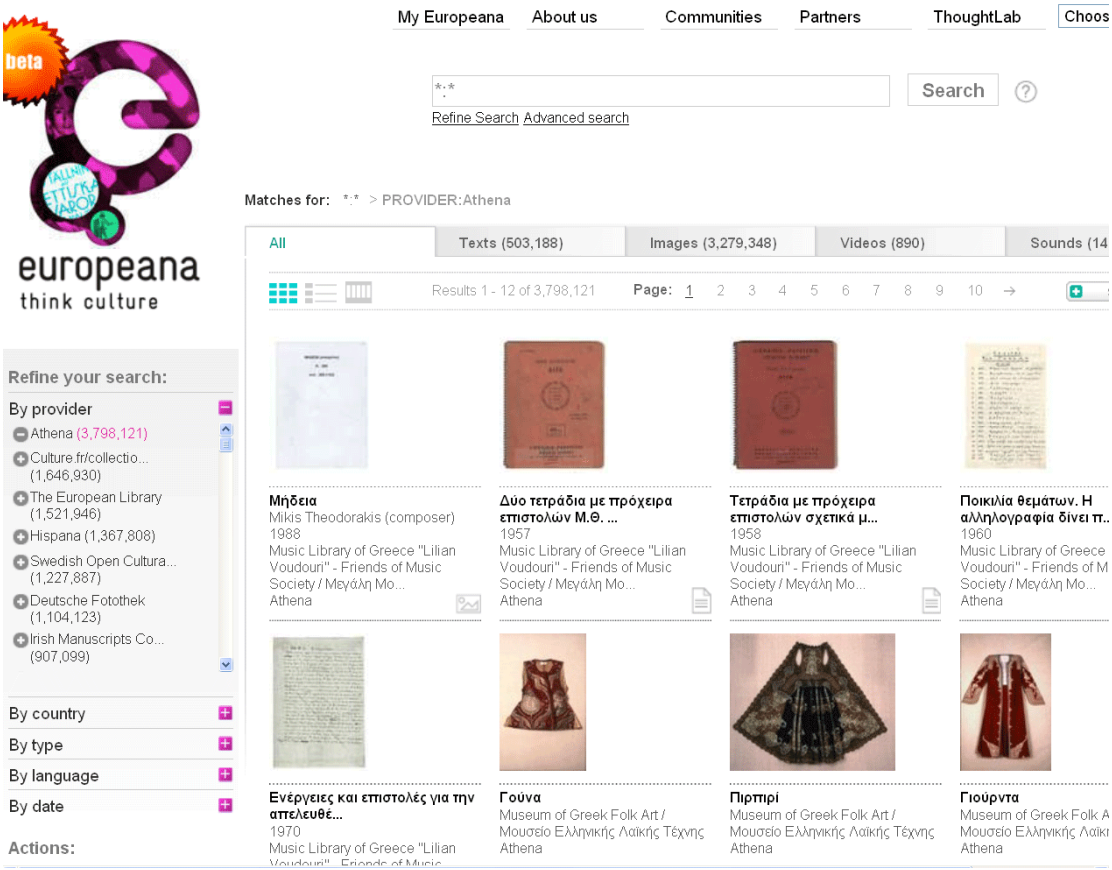



FIG. 2 – ATHENA content in Europeana: 3,798,121 million items.

Obviously there was a substantial back office work carried out by WP5, WP7 and all contributing partners. Managing over 120 cultural institutions belonging to 23 different countries is something serious: plenary conferences, technical meetings, training, daily contact with all partners in the project, over 1,600 emails, plus Skype or telephone conversations. Moreover, the technical solution for the content aggregation (i.e. the ingester) didn't exist at the beginning of the project and it was necessary thinking, building and testing it.

Workflow and aggregation procedures were agreed on the occasion of the first plenary meetings1 and constantly refined as soon as Europeana developed its content strategy.

1 2008-12-10. Roma, Italy. ATHENA kick-off meeting; 2009-02-13. Roma, Italy; . ATHENA second plenary meeting.

## 2. Workflow Step-by-step

This section illustrates how the aggregation process was carried out within the ATHENA project and which cooperation mechanisms among the ATHENA WPs, content providers and Europeana was set up.

### Step 1 – Creation of the ATHENA National Contact Points

The creation of a network referees, the National Contact Points (NCPs), was the first step for managing the consortium; the decision was announced during the second plenary meeting held in Rome on 13th February 2009.
The NCPs main tasks are:

- acting as a bridge between the WP5 coordination and the single national content providers;
- updating the list of the collections to be provided to ATHENA;
- managing the involvement of new content providers;
- disseminating the ATHENA activities at national level.

Figure 3 - ATHENA content coordination map

In this way it was possible for museums to have a contact person representing them in dealing with the ATHENA content coordination team. In turn, the WP5 content management team has the role of interacting with the technical groups both of ATHENA and Europeana that follow the ingestion procedures (FIG. 3).
The list of the National Contact Points is published on the ATHENA web site so to be easily reached by new content providers coming from their own state1. For further information about the role of NCPs see D5.1 'First Report on the network of national Coordination'.

1 http://www.athenaeurope.org/index.php?en/132/national-contact-points

### Step 2 – Investigation of the digital collections

During the period March-April 2009 (months 5 and 6) the National Contact Points were urged to fulfil the first important tasks related to the WPs 3 and 5: the compilation of the standard questionnaire.
The need to further investigate the state of the art of the ATHENA digital collections for the definition of the ingestion plan (WP5 task) and the compilation of a report about the standards applied by the European museums (WP3 task, D3.1), as well as for retrieving information for the other WPs activities (IPR, multilingualism, geo-localisation) were the background of the online survey that was launched.
This action's aims were to have a clear overview on the standards applied by the participant museums and institutions so to build the ATHENA ingestion software, and verify the relevant information for the fields of the ingestion plan (e.g. the subject of the collection, the kind of digital object – text, image, audio, video - the quantity of thumbnails or samples to send to Europeana, if the metadata are aggregated by anyone else, etc.).

All NCPs filled the questionnaire in about their collections. The analysis of the technical standards is gathered in D3.1 'Report on existing standards applied by European museums'.
Answers were online for a very long period and were used to retrieve information for the ingestion plan.

### Step 3 – Ingestion plan

The ingestion plan was agreed with the Europeana representative within the ATHENA consortium with two main goals:

- trying to synchronise the ATHENA and Europeana ingestion and harvesting procedures;
- avoiding overlapping in content provision.

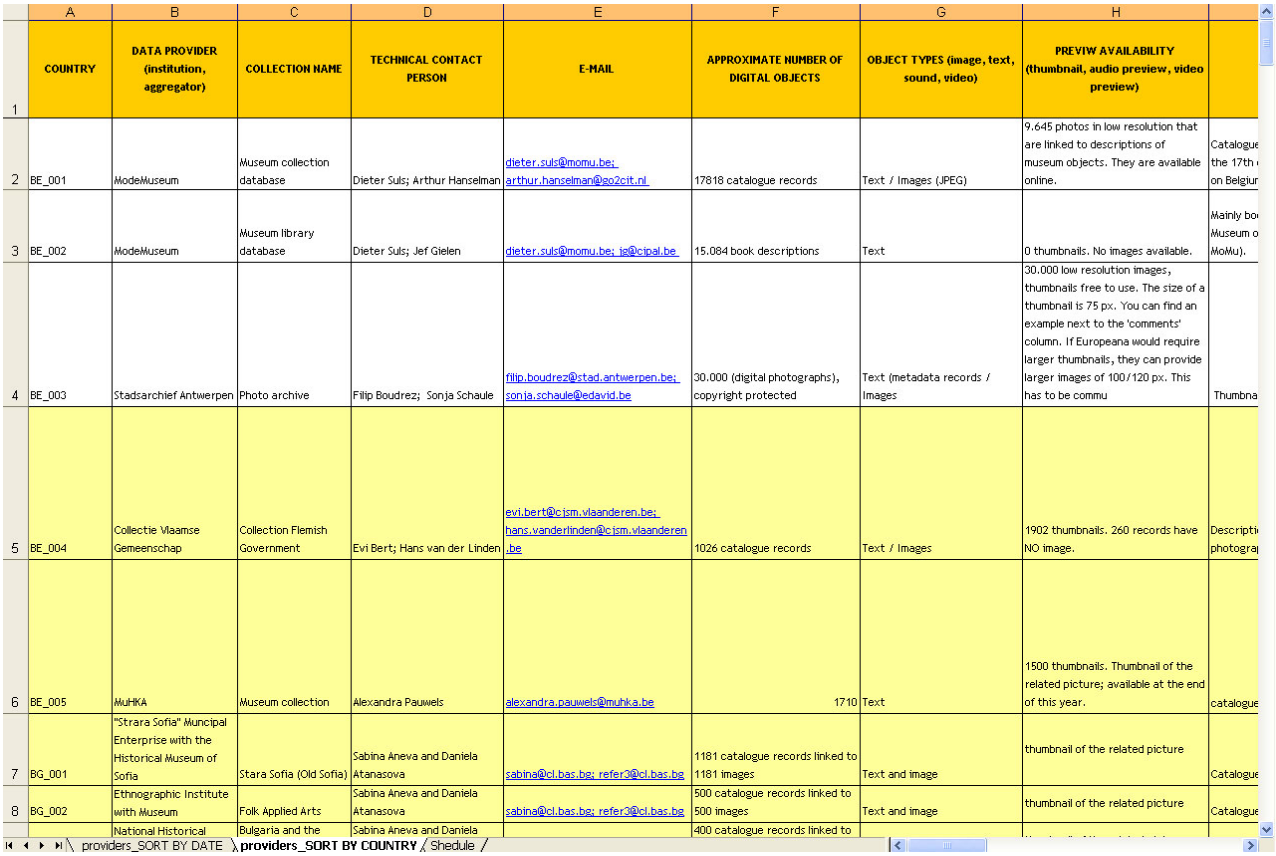
*In fact, little by little that the overall Europeana content strategy was developed (final release in August 2009), it clearly appeared that the fact that the same institutions participate in different EU projects could cause some overlapping in content aggregation and delivery to Europeana (this problem is shared by many of the Europeana family projects).*
*ATHENA WP5 leaders got in touch very soon with the Europeana office to deal this matter and tried to find out the best possible solutions for both projects: it was decided that the ATHENA providers that had already sent their data to Europeana, would have cooperated with their records only to test the ATHENA technical features[1].*

The ingestion plan was structured with the following fields agreed with Europeana (FIG. 4):

- country
- data provider
- collection name
- technical contact person
- email
- approximate amount of digital objects
- object type (image, text, audio, video)
- preview availability
- type of upload (ftp, http, OAI-PMH)

1 D5.3 'Core Content Map for the Recognition of Digital Cultural Heritage Content'.

- used metadata
- technical information in standard questionnaire
- date delivery to ATHENA
- date delivery to Europeana
- comments
- remarks from Europeana
- harvesting authorisation[1]



| | A | B | C | D | E | F | G | H | |
|---|---|---|---|---|---|---|---|---|---|
| 1 | COUNTRY | DATA PROVIDER (institution, aggregator) | COLLECTION NAME | TECHNICAL CONTACT PERSON | E-MAIL | APPROXIMATE NUMBER OF DIGITAL OBJECTS | OBJECT TYPES (image, text, sound, video) | PREVIW AVAILABILITY (thumbnail, audio preview, video preview) | |
| 2 | BE_001 | ModeMuseum | Museum collection database | Dieter Suls; Arthur Hanselman | dieter.suls@momu.be; arthur.hanselman@go2cit.nl | 17818 catalogue records | Text / Images (JPEG) | 9.645 photos in low resolution that are linked to descriptions of museum objects. They are available online. | Catalogue the 17th on Belgiur |
| 3 | BE_002 | ModeMuseum | Museum library database | Dieter Suls; Jef Gielen | dieter.suls@momu.be; jg@cipal.be | 15.084 book descriptions | Text | 0 thumbnails. No images available. | Mainly bo Museum o MoMu). |
| 4 | BE_003 | Stadsarchief Antwerpen | Photo archive | Filip Boudrez; Sonja Schaule | filip.boudrez@stad.antwerpen.be; sonja.schaule@edavid.be | 30.000 (digital photographs), copyright protected | Text (metadata records / Images | 30.000 low resolution images, thumbnails free to use. The size of a thumbnail is 75 px. You can find an example next to the 'comments' column. If Europeana would require larger thumbnails, they can provide larger images of 100/120 px. This has to be commu | Thumbna |
| 5 | BE_004 | Collectie Vlaamse Gemeenschap | Collection Flemish Government | Evi Bert; Hans van der Linden | evi.bert@cjsm.vlaanderen.be; hans.vanderlinden@cjsm.vlaanderen.be | 1026 catalogue records | Text / Images | 1902 thumbnails. 260 records have NO image. | Descripti photogra |
| 6 | BE_005 | MuHKA | Museum collection | Alexandra Pauwels | alexandra.pauwels@muhka.be | 1710 | Text | 1500 thumbnails. Thumbnail of the related picture; available at the end of this year. | catalogue |
| 7 | BG_001 | "Strara Sofia" Muncipal Enterprise with the Historical Museum of Sofia | Stara Sofia (Old Sofia) | Sabina Aneva and Daniela Atanasova | sabina@cl.bas.bg; refer3@cl.bas.bg | 1181 catalogue records linked to 1181 images | Text and image | thumbnail of the related picture | Catalogue |
| 8 | BG_002 | Ethnographic Institute with Museum | Folk Applied Arts | Sabina Aneva and Daniela Atanasova | sabina@cl.bas.bg; refer3@cl.bas.bg | 500 catalogue records linked to 500 images | Text and image | thumbnail of the related picture | Catalogue |
| | | National Historical | Bulgaria and the | Sabina Aneva and Daniela | | 400 catalogue records linked to | | | |

providers_SORT BY DATE | providers_SORT BY COUNTRY | Shedule

Figure 4 - ingestion plan snapshot

The core information of the ingestion plan is:

- approximate amount of digital objects > to fulfil the DoW mission
- object type (image, text, audio, video) > to be sure that the digital format is compliant to the Europeana requirements
- preview availability > Europeana wants previews: if they don't exist, they must be created
- used metadata > to keep under control interoperability
- remarks from Europeana > notes from the Europeana ingestion team (e.g.part of the collection doesn't meet the Europeana mandatory features

The ingestion plan is a working and evolving document in constant evolution; periodically it is uploaded into the reserved area of the ATHENA website and is intended as Annex I of this deliverable.

1 See paragraph 'Europeana agreements' pp. 17-18.

**THE INGESTION PLAN AT A GLANCE**[1]

**Quantity of content to be uploaded in the ATHENA ingestion tool**: 4,6 M about

**Number of digital collections**: 225 (+11 to be confirmed)

**Providers**: 122 (+5)

**Countries**: 23 (including Israel and Russia)

**Subjects**: fine arts (paintings, installations, etc.), architecture & landscape, dress & textiles, photographs, postcards, folk applied arts, folk music, manuscripts (including the 'Shrine of the Book'), magazines and books of various ages, numismatics, archaeological artefacts (Roman, Hellenistic, Byzantine, prehistoric), autographs, anthropology collections, minerals, literary works, sculptures, ancient glass, classical archaeology, design, prints and drawings, maps, technical machines, industrial archaeology, archival material

### Step 4 - The ATHENA ingester

The National Technical University of Athens (NTUA), leader of WP7 'Development of plug-in to be integrated within Europeana' developed specific software to facilitate the mapping and the ingestion of content supplied by the project partners, and the semantic interoperability in the Europeana context[2].

This metadata ingestion service, commonly called 'ingester', enables:

- the data uploading and mapping to LIDO (by contributing partners);
- the transformations of the metadata records into LIDO records and the aggregation in the ATHENA repository;
- transformation of stored data into the ESE version and transmission to the Europeana ingestion office via OAI-PMH.

The metadata ingestion service needed to be built up and content providers started testing it only at the end of 2009.

The ATHENA ingestion service implemented LIDO v.0.93, the XML schema elaborated for delivering different museum metadata to cultural heritage repositories made up of a comprehensive format including as much information as possible to avoid any loss of granularity. The ingester is structured around LIDO and enables the transformation into the

1 Details on content provision are given in D5.3 'Core content map for the recognition of digital cultural heritage content'.

2 See D7.4 'Report on the integration of the plug-in with the Europeana portal', pp. 15-33. The ATHENA tool can be reached at the following URL: http://athena.image.ntua.gr/athena (reserved to the ATHENA content providers for the ingestion of their contents).

3 D3.3 'Specific for conversion tools'. Cfr. E. Coburn, R. Light, G. McKenna, R. Stein, A. Vitzthum, LIDO - Lightweight Information Describing Objects Version 1.0, November 2010: 'Organizations need to provide information on their objects to many portals including thematic, cross domain, regional, national and international, and web applications. The difficulty is that the object information is in the providers' own collections management systems and cataloguing databases. Each of these has potentially a different metadata format. This means that it is both time-consuming and costly to integrate information from all those organizations wanting to participate. To overcome this situation LIDO has been developed' (www.lido-schema.org/schema/v1.0/lido-v1.0-specification.pdf).

Europeana Semantic Element (ESE). The ingester also embeds the Europeana Content Checker for validating content before the publication1.
The ATHENA ingestion tool is able to aggregate content from various cultural fields (not only from museums) and structured with different schemas.

What a content provider must do to upload data?

- create an account through the 'Register' page
- edit and save the mapping to LIDO
- upload the XML files (through http, FTP, OAI) (FIG. 5)
- overview the imports
- visualise previews in XML LIDO and XML ESE
- visualise Europeana preview
- if the result is acceptable, press the publishing button and wait for the Europeana harvesting

The whole workflow is subordinated to the subscription of the ATHENA licence that regulates the content provision (see further paragraph 'The Europeana agreements' p. 17).

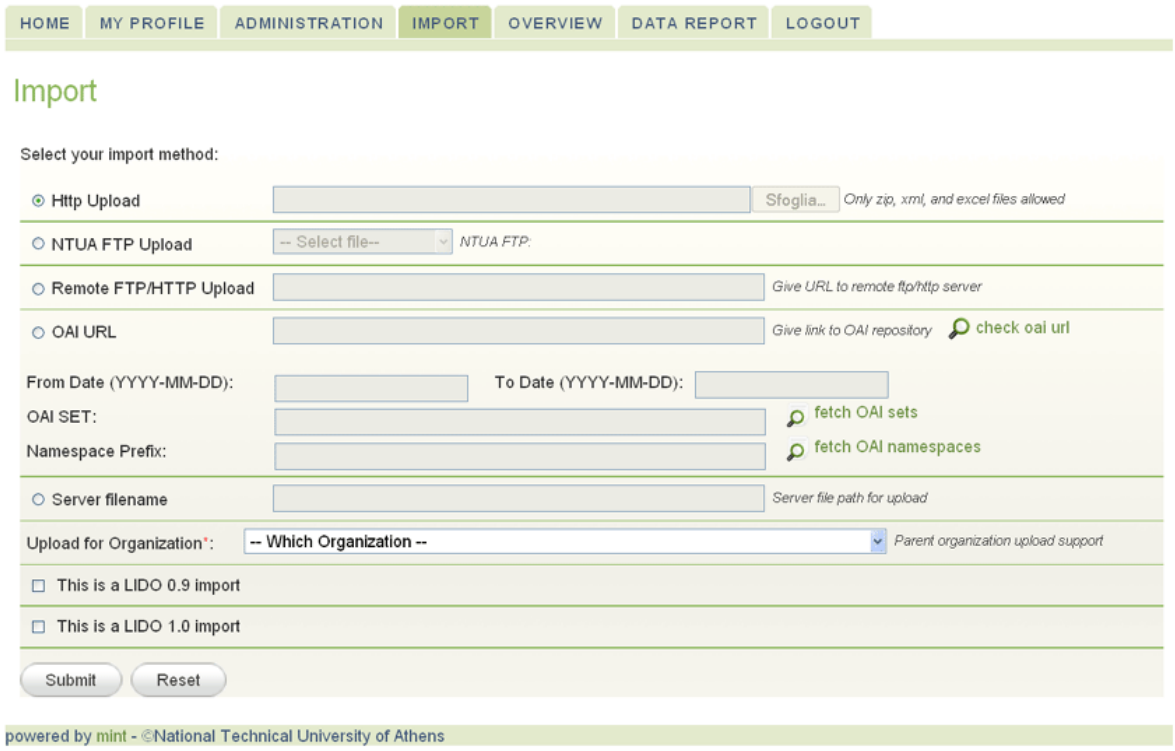



Figure 5 – ATHENA ingester – import interface

1 Europeana Content Checker User Guide. http://www.google.com/url?sa=t&source=web&cd=1&ved=0CBgQFjAA&url=http%3A%2F%2Fversion1.europeana.eu%2Fc%2Fdocument_library%2Fget_file%3Fuuid%3Dd9ca0106-affb-4a38-83e9-e886289dd0d9%26groupId%3D10602&rct=j&q=europeana%20content%20checker&ei=o3fwTeD2FcaAOoDA0LID&usg=AFQjCNHumQfUlBbaKdZSreFXNmOx0ah4Rw&cad=rja

In October 2010 the 'Guidelines for the publication on Europeana' were issued and communicated all partners.

**AGGREGATION PROCEDURE**

- First of all, the partners who have not done yet, MUST immediately register in the ATHENA ingestion system (http://athena.image.ntua.gr/athena). For any question about the procedure, please e-mail to G. (CCing V. and M.).
- As soon as possible, every partner SHOULD UPLOAD AND MAP TO LIDO a sample of data that covers all the elements in XML file (even a small set of records is sufficient for the test).
- When the data sample has been uploaded, the partner SHOULD INFORM V. and G. (CCing M.). For any problem with the uploading/mapping, please send an e-mail to G. (CCing V. and M.) who will review data in order to make the partner able to proceed with the final uploading.
- Before the publishing procedure, ASK FOR FINAL AUTHORIZATION to Gordon (CCing V. and M.).
- After the authorization, the partner WILL PRESS THE PUBLISHING BUTTON (please inform V., G. and M.) in order to activate the harvesting process by Europeana.
- If there are changes, respect to the previous mapping, the partner MUST ASK AGAIN FOR THE VALIDATION. If there are no changes, partners CAN press the publishing button but they MUST inform V., G. and M. of the new situation.
- It is important that if the partners have collections not yet ready to be uploaded, SHOULD programme periodic upload in order not to overcrowd the system and the procedure in the last period of the project.

**OTHER GUIDELINES**

- Partners SHOULD supply a link to a thumbnail preview which is preferably the same width or slightly larger than the requirements of Europeana (currently 200 pixel wide). Partners SHOULD NOT supply a link to a very large image (i.e. much wider than 200 pixels). If the thumbnail is larger than the requirements, Europeana WILL REDUCE it. If smaller, Europeana WILL NOT RESIZE it (It is recommended that content providers look into possibilities to supply bigger size source images for image caching if possible).
- Partners SHOULD NOT supply a link to a 'dummy' thumbnail that indicates that there is no image or that no image can be displayed for IPR reasons. If this in not easily possible, partners SHOULD supply the URL of the 'dummy' thumbnail to Gordon.
- Europeana WILL NOT ACCEPT records without showing a digital object (image, etc.) on the partner's site. However, these records CAN BE UPLOADED on the ATHENA tool.

The complete set of guidelines on how to map, ingest and interact with WP5 is published on the ATHENA web site: the content providers can find there the ingester tutorial, introduction to LIDO and tools for mapping (including a mapping worksheet), the videos of the training meetings[1].

1 http://www.athenaeurope.org/index.php?en/159/training

Shorter guidelines on how to deliver digital content to Europeana through ATHENA devoted to new cultural institutions are also published in D2.3 'Set of instruments to support newcomers to join'.

The task of mapping data to LIDO and using the ATHENA ingester is a winning choice in the light of the forthcoming Europeana changes. The switch from the Europeana Semantic Elements (ESE) to the new Europeana Data Model (EDM), conceived for preserving the original information and the richness of relations between the objects, will be soon realised. The complexity of LIDO is nearer to EDM rather than ESE; furthermore, the ATHENA ingester will store for next 3 years1 the ATHENA data already mapped to LIDO: once that Europeana will shift to EDM, it won't be necessary that the ATHENA providers re-map and re-upload all data to see them published in Europeana, but only that a mapping LIDO-EDM be provided.

The ingester is a useful tool not only for content management and aggregation carried out by providers, but also for the monitoring task that WP5 has to perform. The following pictures (FIG. 6 and 7) illustrate how it is easy having a general overview of the uploaded content.

| Overall | |
|---|---|
| Registered users: | 164 |
| Registered Organizations: | 152 |
| Total Items: | 4647832 |
| Total LIDO Items: | 4529474 |
| Total ESE Items: | 4004949 |
| Countries contributing: | Belgium, Bulgaria, Cyprus, Czech Rep., Estonia, Finland, France, Germany, Greece, Hungary, Israel, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Romania, Russia, Slovakia, Slovenia, Sweden, United Kingdom |
| Countries with no data: | Austria, Denmark, Ireland, Portugal, Spain, Switzerland, Europe, International |

* Reporting on ESE items comes from the ATHENA OAI repository where data is available for harvesting by Europeana



FIG.6 - Overall amount of uploaded files into the ATHENA ingester at 26 April 2011: 4,647,832 uploaded metadata, 4,529,474 of which transformed into LIDO and 4,004,949 into ESE. On the right, a spotlight on the ingestion of a content provider (the Lithuanian Art Museum).

The ingestion software makes it possible to monitor in real time the position of contributing countries, the state of the art of loading data, the total number, the number of data mapped in LIDO and its processing in ESE, that is the data that are available for harvesting by Europeana and then ready for publication on line.

Through the overviews of the software it is possible to see both items imported by the various institutions and the uploaders. It is also possible to check the metadata transformed in ESE, and then ready for publication in Europeana.

1 For the duration of Linked Heritage project at least (see D1.5 'Sustainability Concept').

In fact, through a button representing the logo of Europeana the content providers have the possibility to make their data at disposal of the harvesting and final publication in Europeana. The content provider has the opportunity to choose which data publish or not.
The following image (FIG. 7) the state the state of publication of Bildarchiv Foto marburg in the ATHENA tool: some datasets are ready to be harvested by Europeana (they have the Europeana logo beside) and others are only tests so they have not been published for the Europeana harvesting.

FIG. 7 - XML sets report. The Europeana logo on the left shows that data have been published.

Content providers can also check their records in XML LIDO (FIG. 8), XML ESE (FIG. 9), and as they may be shown in Europeana (FIG. 10), so to eventually modify them before the effective publication.

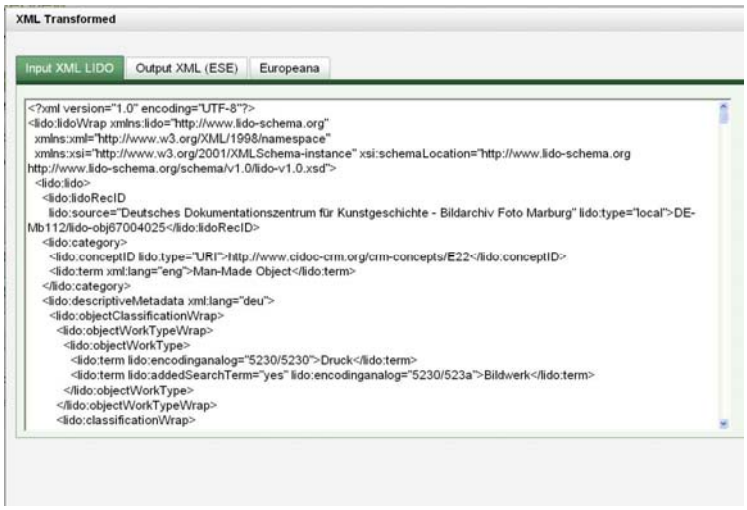



FIG. 8 – Input XML LIDO.

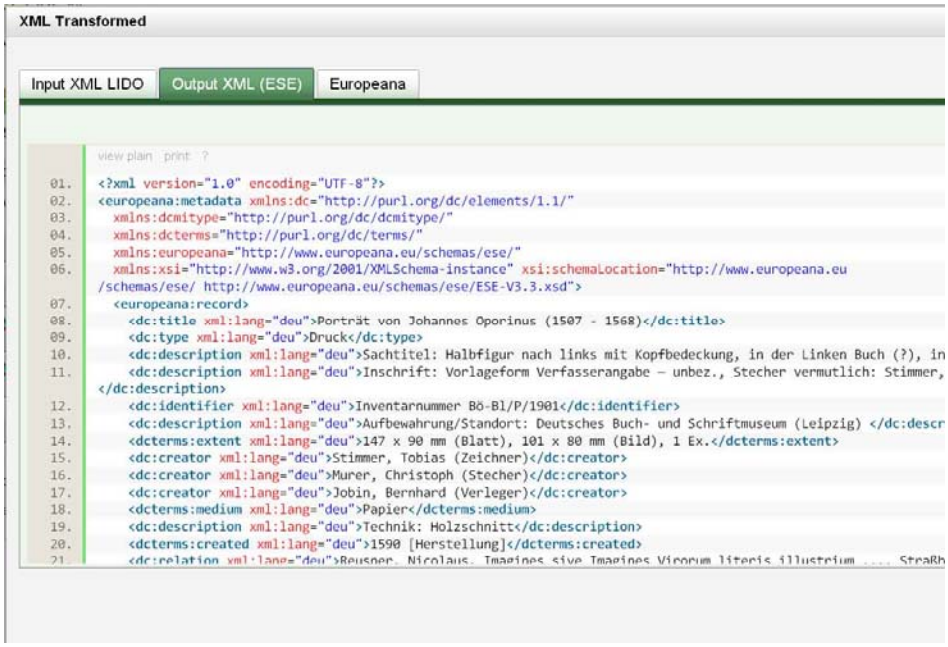



FIG. 9 – Output XML ESE.

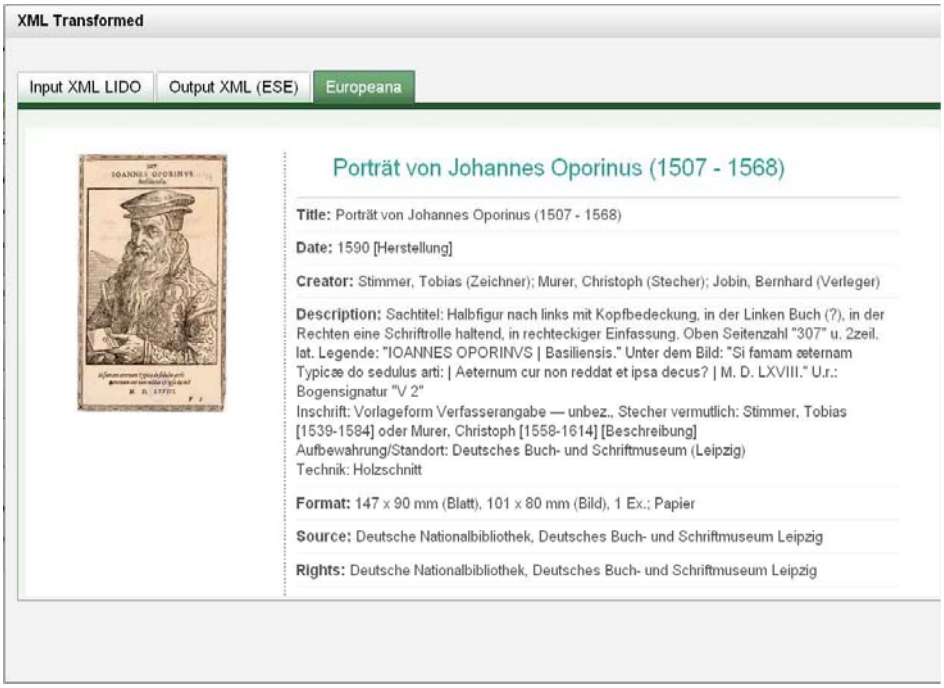



FIG. 10 - Europeana preview.

The result in Europeana: 3,798,121 items provided by the Athena Project are currently online. With the finalization of the Europeana publication will be available on line more than 4 M data, which will increase with the participation of new cultural institutions, and new data in the next future.

### Step 5 – Training and training material

Face-to-face lessons are the best way to train content providers on the mapping to LIDO and the use of the ingester; training workshops were organised in cooperation with WP3 and 7 in:

- Athens (November 2009)
- Rome (January 2010)
- Berlin (January 2010)
- Ljubljana (June 2010)
- Athens (February 2011)

Different periods and places allowed almost all content providers to participate. In Rome and Ljubljana it was also possible to show live via streaming the complete workshops; the videos are published on the ATHENA website in the section of training material, together with all the other tools and guidelines to support content providers in the use of LIDO and the ingester1:

- The ATHENA LIDO Mapping Worksheet
- An ATHENA Ingester Tutorial
- A set of Basic Rules for Mapping
- A set of Guidelines on the Use of the ATHENA Ingester
- A set of Guidelines for Publication on Europeana

1 http://www.athenaeurope.org/index.php?en/159/training.

- Some Clarifications concerning the ATHENA server
- Step by Step IPR guide[1]

**Step 6 – ATHENA help desk**

As soon as the ingestion phase started up, a helpdesk was created to support the content providers (athena-helpdesk@amitie.it). Thanks to the cooperation among WPs 3, 5, and 7, any kind of critical issue could be shared and discussed: leaders and experts of these groups answered questions concerning mapping (WP3), technical procedures (WP7), and collection management (WP5), as well as the relationship with Europeana. Providers could join the group upon subscription[2].
The service has proved to be an important support both in providing solutions both in sharing problems and experiences during all the ingestion process.

1 http://www.athenaeurope.org/index.php?en/192/step-by-step-ipr-guide

2 Registration form at http://www.athenaeurope.org/index.php?en/157/registration-form-athena-technical-help-desk-mailing-list.

## 3. Criticisms

It is useless to hide the fact that a process of aggregation of metadata as a whole involved many difficulties.

First, there were some technical criticisms:

- the system managing the digital collections needed to be re-engineered: this caused many delays;
- as every software, sometimes the systems managing the partner digital collections could brake down and could be was necessary too much time to recover it;
- software bugs (both of the ATHENA ingester and of Europeana)
- metadata were not associated to digital objects (text, image, audio and video) like catalogue library records: they are rejected by Europeana. When the ATHENA proposal was written, Europeana didn't have yet a clear content strategy and it was not acknowledged that metadata only were not enough;
- rich metadata but low quality images (this, for museum objects, is a terrible criticisms!)
- customisation of metadata standards: difficult interoperability.

Then, management difficulties arose:

- the collections put at ATHENA's disposal were partially digitised: at the time of the proposal the content providers assumed that the digitisation process would have been completed for the beginning of the project;
- same collections in ATHENA, Europeana and related projects (which rules for directing the traffic?);
- the collection manager changed or moved to another office: there was a vacancy for some time and then the new manager needed to be trained and informed over again;
- copyright issues: some institutions didn't have the complete ownership of the rights on the digital collections, in particular on those resulting from a joint work carried out with other cultural bodies. This caused delays for the right clearance or the cancellation of the collection;
- synchronisation of the activities within the ATHENA consortium and then with Europeana. Metadata face 2 or 3 different aggregation steps1 from the providing museum to Europeana: a delay in one of these steps breaks the plane of aggregation with a snowball effect.

And finally aggregation problems:

- mapping errors;
- thumbnail blurring;
- problems with image caching by Europeana;
- possible lack of visibility of content providers in Europeana (if metadata are provided via an aggregator).

All these critical issues generated a mass of hundreds of conversations with the ATHENA helpdesk and the Europeana ingestion team in an attempt at solving them:

> *'G. asked me to fix dcterms:title and change it to dc:title, map dc:source to europeana:dataProvider, leave dc:description as it is. I cannot change the display*

1 Museum > ATHENA > Europeana [2 steps] OR Museum > aggregator (national, thematic, regional) > ATHENA > Europeana [3 steps].

*order of dc:extent. Dc;Type was another issue I now mapped it to dc:subject because this fits better. I tried to differenciate between dc:type and dc:subject but currently this information is all mapped to dc:type in the input file and there is no pattern that I could filter for. Not all items contain a true dc:type as first element in their metadata'.*

Problem solving was not always easy and of course bad digitisations couldn't be changes; nonetheless, the final satisfying results of the projects demonstrated a strong cohesion of the activity leaders in managing as a whole both technical and management criticisms.
Technical criticisms were overcome thanks to the refinement of the mapping procedures (by providers) or the ingester (by NTUA); by the other hand, the Europeana ingestion team was very active in trying to manage the criticisms due to their system. WP5 tried to solve problems connected the digital collections that were put at the project's disposal.

### Collection criticisms

Criticisms concerning collections were illustrated in details during the first ATHENA review and in the management deliverables1. This mainly concerned collection overlappings.
The overlapping of content i.e. same collections in ATHENA and Europeana or other related projects were faced in agreement with Europeana as follows:

- in case of same collections in ATHENA and Europeana, the collections could be provided to Europeana directly and eventually to ATHENA only to test the ingester and LIDO (as, for instance, for Collections – the French culture portal – or English Heritage)
- the provider managed to give both to Europeana and ATHENA different digital collections (e.g. Ministry of Education and Culture and Department of Antiquities from Cyprus, CIMEC, etc.);
- the provider goes to the project that better fits the content: e.g. archives partners of the ATHENA consortium agreed to reach Europeana through APEnet, the archive aggregator, or AV archives passed to EFG (this is for Greek Film Archive).

ATHENA WP5 worked in close cooperation with the Europeana office to avoid other overlapping and to fulfil what was written in the Description of Work.

### The Europeana agreements

The Europeana agreements as they are now are not a critical issue since they have been acepted by the ATHENA content providers; nonetheless, when they were published as a draft they created an immense uproar across the ATHENA community and some content providers weren't favourable anymore to supply Europeana with their data. This fact took time and work to be overcome.

In 2009 Europeana issued the licensing framework, the agreements for providers and for aggregators, to regulate the publication and use of the metadata. Before the final release the agreements circulated among the Europeana projects' communities for feedbacks. Those ones

1 See in particular D1.2 'First annual report' pp. 28-30.

expressed by the ATHENA consortium were collected into a document[1] (issued in January 2010) that was sent to Europeana whose general feeling can be resumed as follows:

> *'We think that Europeana is asking for way too many rights from the museums' (W. S., Bibliotheksservice-Zentrum Baden-Württemberg).*

The main concern of the European museums was that the commercial use of the metadata must be explicitly excluded. For the museum communities the metadata of museum objects can be considered as small essays and moral rights lies on them.
The final version of the agreements took care of excluding the commercial purpose and, as a consequence, they were accepted by the ATHENA consortium. However, partners asked the project coordinator to regulate the data transfer to Europeana with written formal authorisation referring to the Europeana Licence Agreement[2]; this passage was not expressed in the Description of Work or requested by Europeana but was helpful to clear duties and rights of all parties.

By the way at the end 2010 Europeana published the first draft of the Data Exchange Agreement that will substitute the previous licences. The ATHENA consortium contributed to the development of this new regulating framework participating at the expert meetings and organising in April 2011 a workshop in Brussels on Europeana's Data Exchange Agreement in order to discuss all the arisen concerns[3].
This new agreements requires that all resources are made available to Europeana with a CC0 Public Domain Dedication licence. The main purpose for this major change is the possibility to publish Europeana data as Linked Open Data (LOD); according to Europeana it is necessary to remove the non-commercial close for working with LOD and this frightened some ATHENA partners that instead would like to be sure that their resources are used only for discovery purposes within Europeana.

The switch toward the Data Exchange Agreement still didn't happen but it is now clear that the direction that Europeana will take is to include the possibility to reuse data as LOD. After the end of the ATHENA project partners will be contacted by Europeana and invited to sign the new licence[4] but currently some of them is still doubtful.

1 'Europeana Data Aggregator Agreement: ATHENA comments'.

2 See details in D1.3 'Second Periodic Report' pp. 12-13.

3 'Europeana Consultation on New Data Provider and Aggregator Agreement: ATHENA partners' comments', 15 january 2011. See D 1.5 'Final report' for detail on the issue.

4 See D1.5 'Sustainability concept'.

# 4. Lesson Learnt

Setting up the ingestion tool and the ingestion procedures was instructive for the ATHENA stakeholders. Many are the lessons learnt, someone evident, and others less:

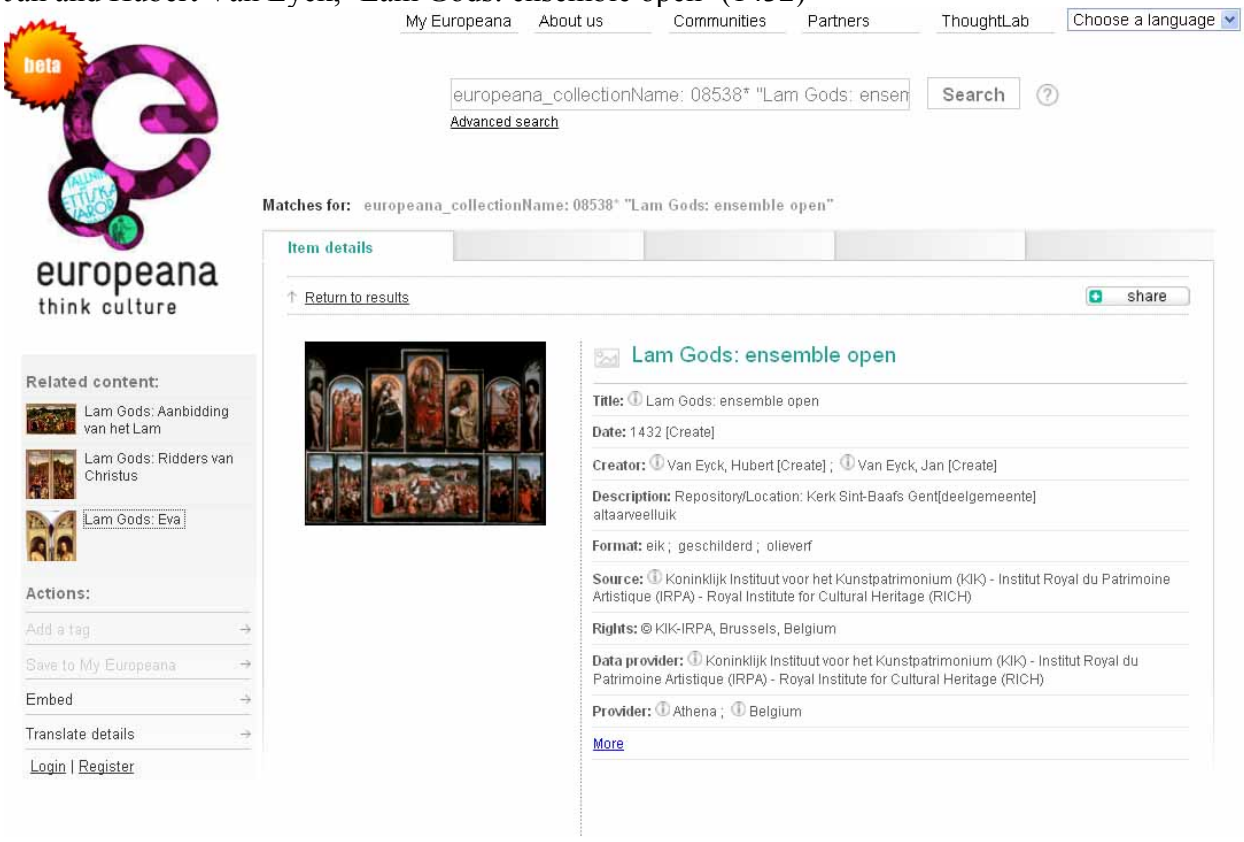
- every content aggregation process, including the Europeana one, is a constantly evolving practice and the technical aspects are relatively relevant: timing and human factor ('the metadata expert is going on holiday; we should postpone the image caching' -!-) are equally relevant;
- the importance of respecting technical guidelines and standards in the digitisation process, so to build interoperable systems. Unfortunately the lack of application of standards or – which is worth – the customisation of standards (that become, de facto, non-standards) is widespread yet.
- training on digitisation, at both basic and professional levels, for museums experts in particular. Domains other than museum one, library for instance, are more used to apply standards for the cataloguing, creation and management of digital resources;
- expert networking for problem solution: the example of the ATHENA helpdesk was an illuminating example;
- the 'human interoperability', i.e. the will to cooperate between ICT and cultural heritage experts is the conditio sine qua non for the realisation of huge initiative like ATHENA.

# 5. Masterpieces

ATHENA is not only the largest Europeana aggregators, but also a provider of masterpieces. Some of them – real world heritage - were already introduced in deliverable 5.3: from Giotto's painting to the Parthenon's marbles, from Damien Hirst works to Alvar Aalto drawings, etc1.
Some of these masterpieces were showed during the ATHENA final conference; afterwards many other content providers highlighted the main pieces of the collections they provided (or are providing to Europeana). Some examples follow; captions have been given by collection managers to motivate their choices.

## BELGIUM

Royal Institute for Cultural Heritage (KIK-IRPA) - Brussels
Jan and Hubert Van Eyck, 'Lam Gods: ensemble open' (1432)

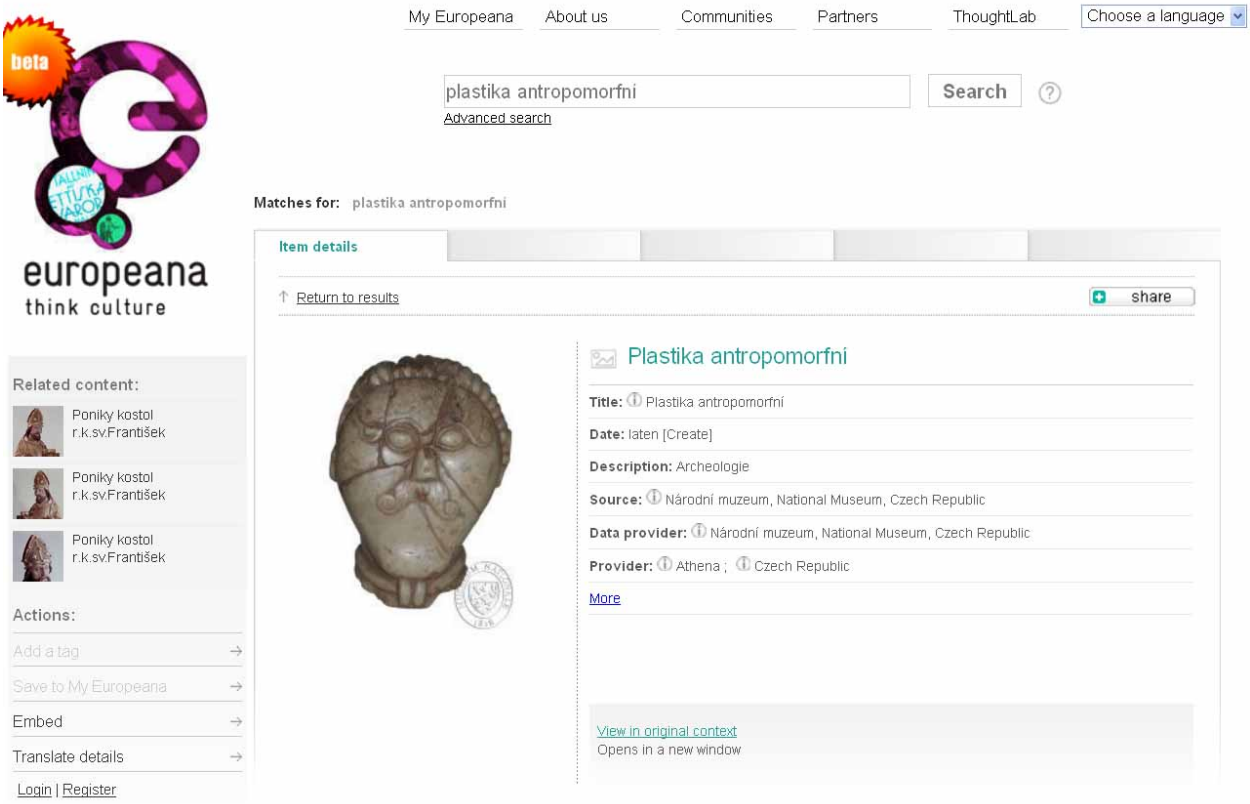



One of the most important pieces of the KIK-IRPA collection in Europeana is the "Ghent Altarpiece" ("Lam Gods"/ "Agneau Mystique") painted by the brothers Jan and Hubert Van Eyck in 1432 and a world masterpiece of the Flemish primitives. It is in the Saint-Bavon cathedral in Ghent (Belgium).

1 D5.3 'Core Content Map for the Recognition of Digital Cultural Heritage Content' pp. 15-17.

## CZECH REPUBLIC

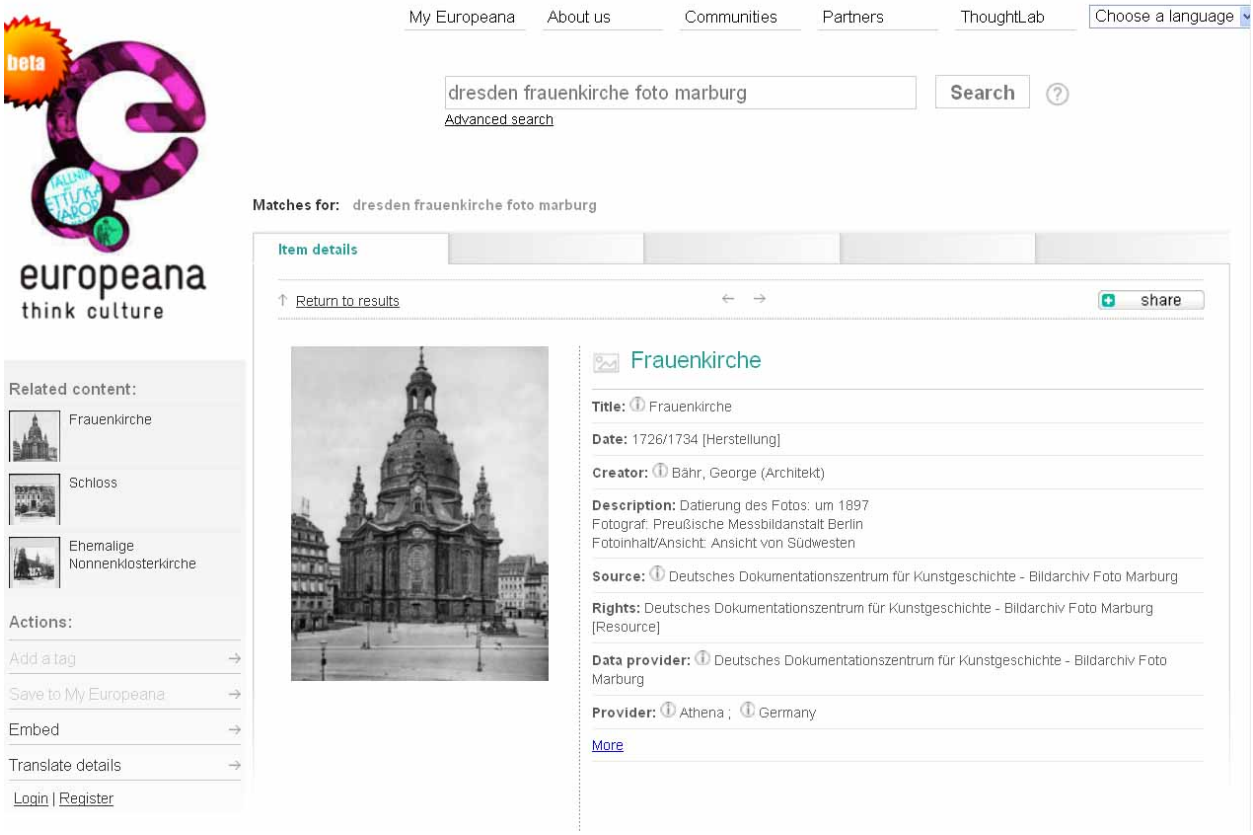
Národní museum - Prague




The most important item for tor the National Museum is the stone head wearing a torc from Mšecké Žehrovice (late La Tène culture). This is masterpiece is symbol for Celtic habitation in Bohemia and it is mentioned in all books about Celts.

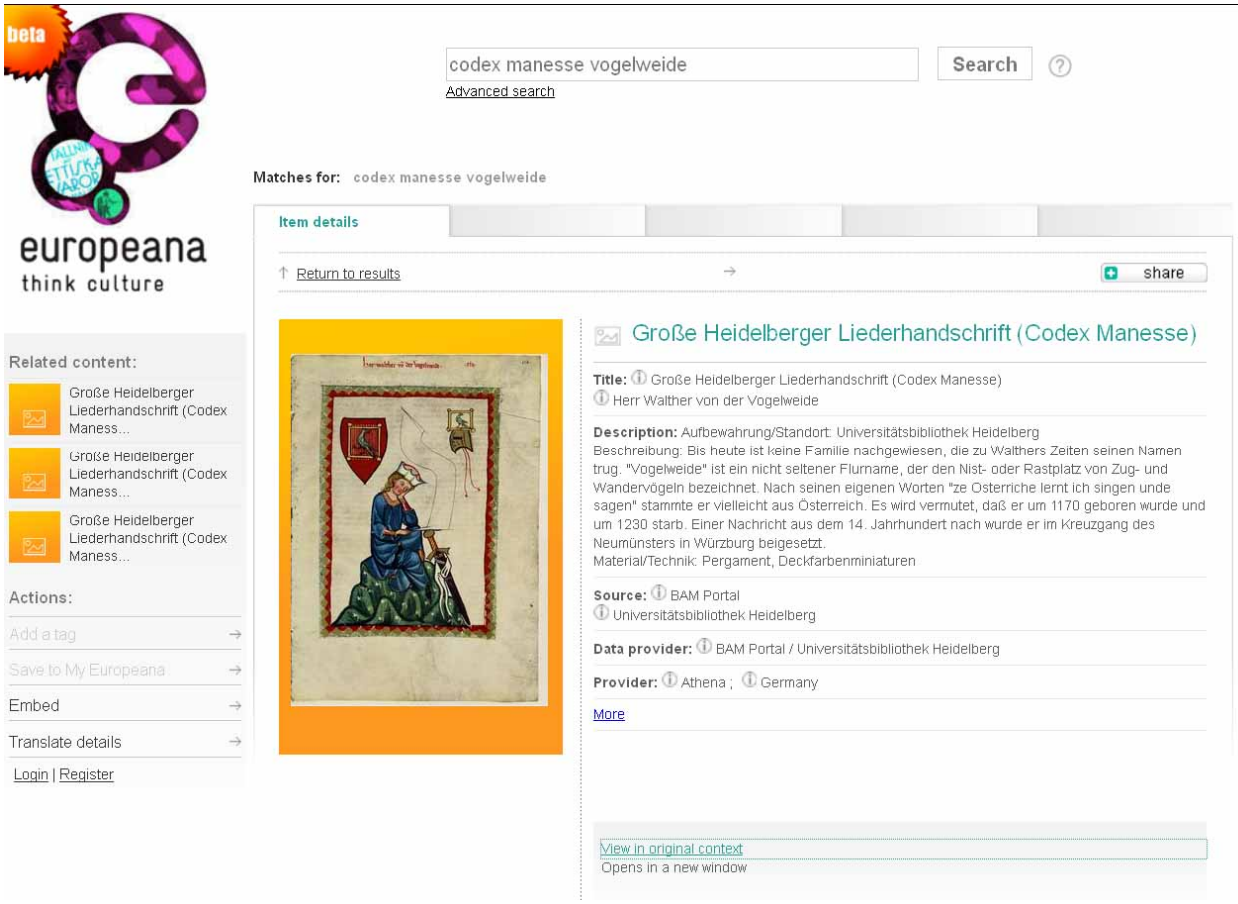
## GERMANY

Bildarchiv Foto Marburg - Marburg



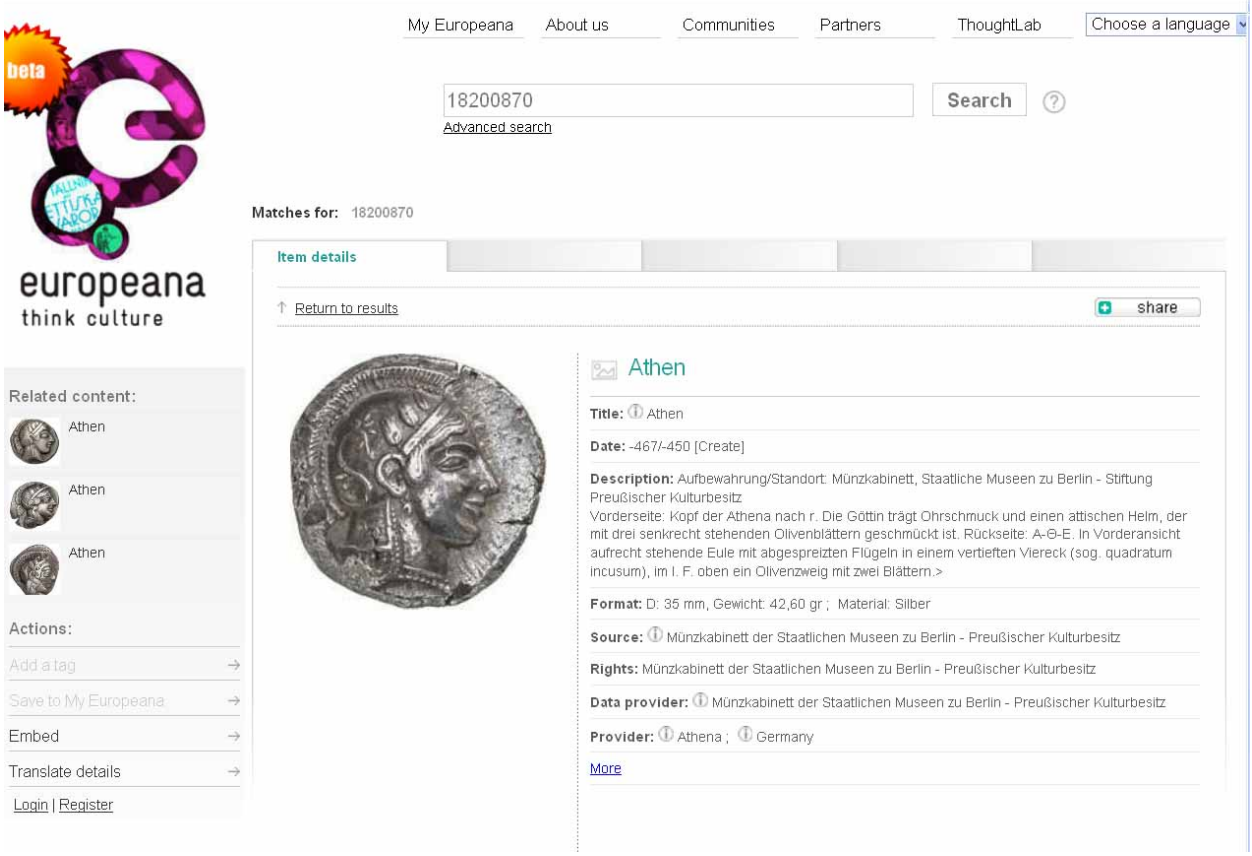

The Lutheran church "Frauenkirche" in Dresden, Germany was built in the 18th century, and was destroyed in the firebombing of Dresden during World War II. It has been reconstructed as a landmark symbol of reconciliation. The photographic view of the church was taken in 1897.

Universitätsbibliothek Heidelberg - Bibliotheca Palatina Baden-Wurttemberg - Heidelberg Codex Manesse



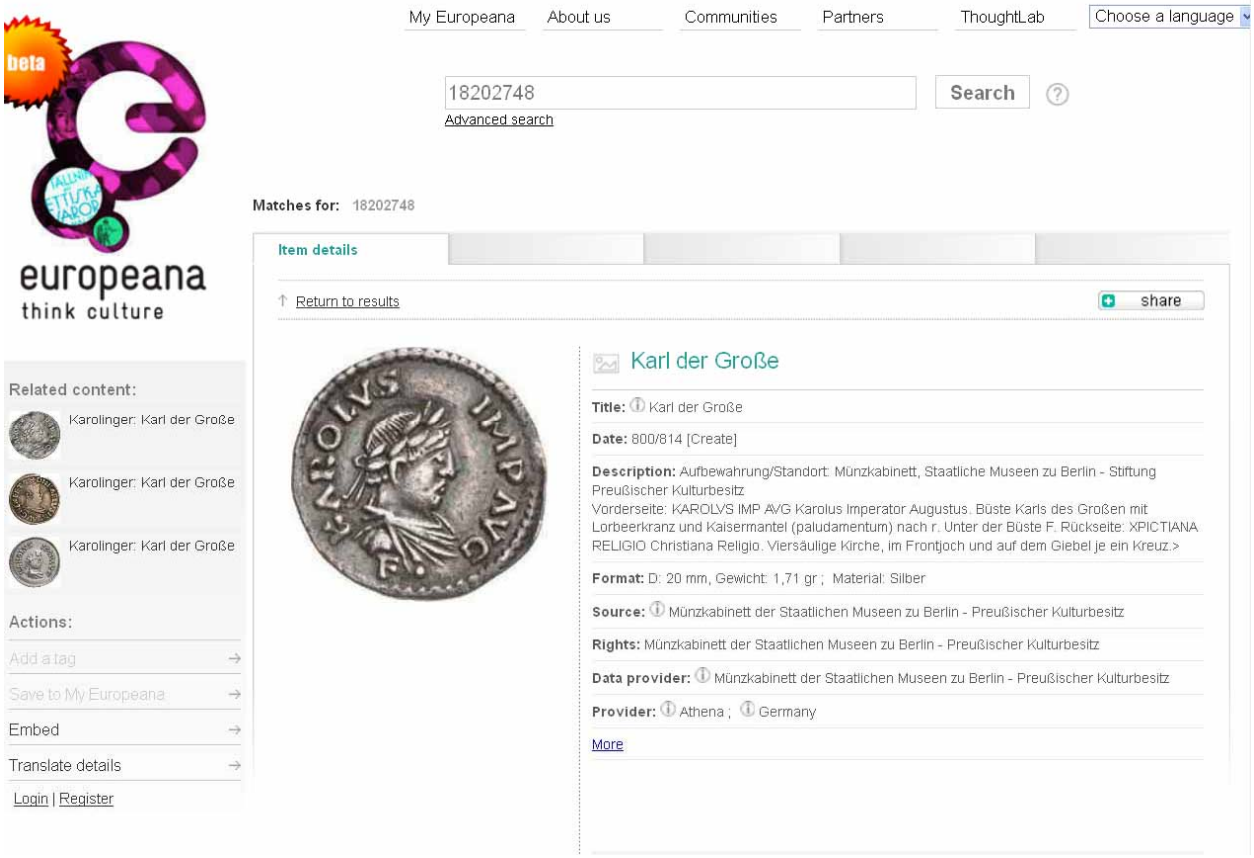

The exquisitely illuminated large format manuscript is one of the most famous medieval manuscripts. It was created at the beginning of the 14th century in Zurich presumably due to the initiative of Johann and Rüdiger Manesse. The manuscript unites the entire range of Middle High German lyric, in its diversity of form and genre. The oldest texts in the Codex Manesse reach back as far as the mid 12th century, and many of the poems are uniquely documented here: The Codex Manesse is, thus, one of the key testimonies for the literature and culture of the Staufer dynasty.

Münzkabinett SMB - Berlin
Decadrachm, Athens 467-450 BC



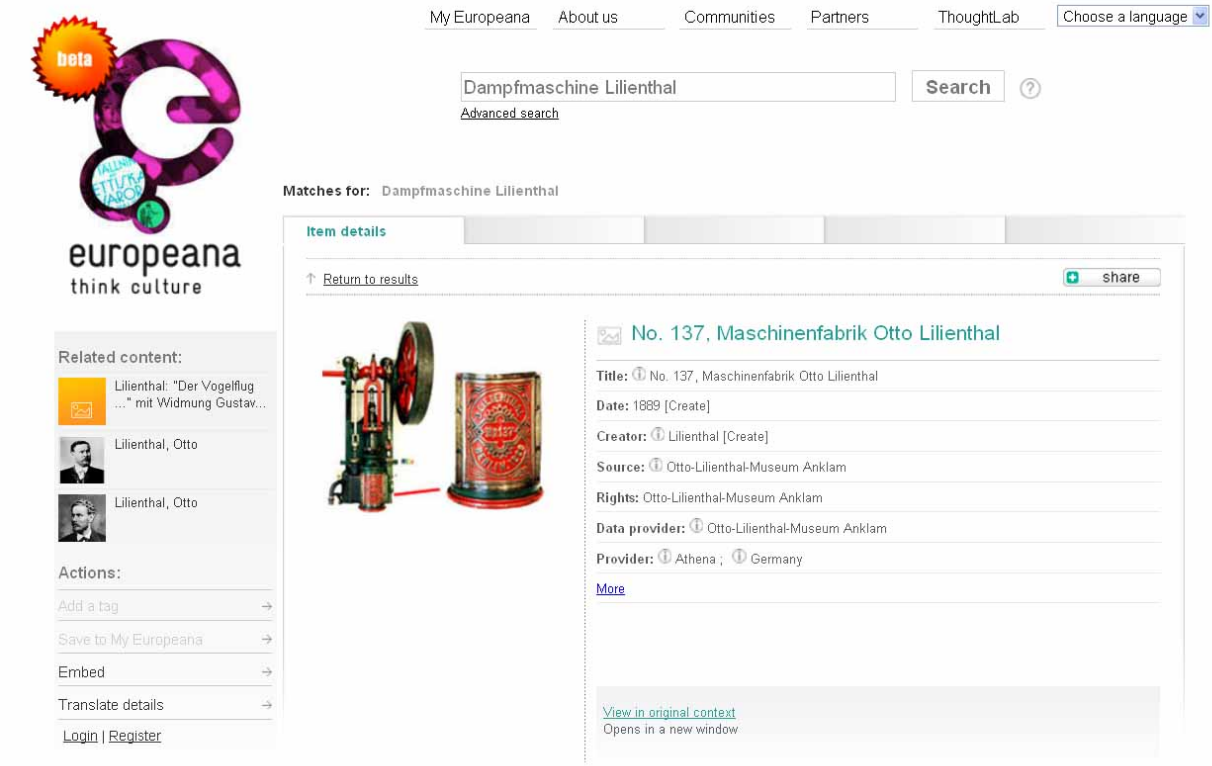

This particular coin is a decadrachm, a silver coin worth 10 drachmas. Tetradrachms (i.e. coins to four drachmas) were common in Athens. The decadrachm towered over all other Athenian coins in size and weight of this coin - one of the best preserved - comes from the almost 11,000 coins comprehensive collection of the ambassador at the Sublime Porte (the Ottoman Empire) and in Athens, the earl  Anton Prokop von Osten (1795 -1876), which was acquired in 1875 by coin collection.

Denar of Charlemagne (800-814)



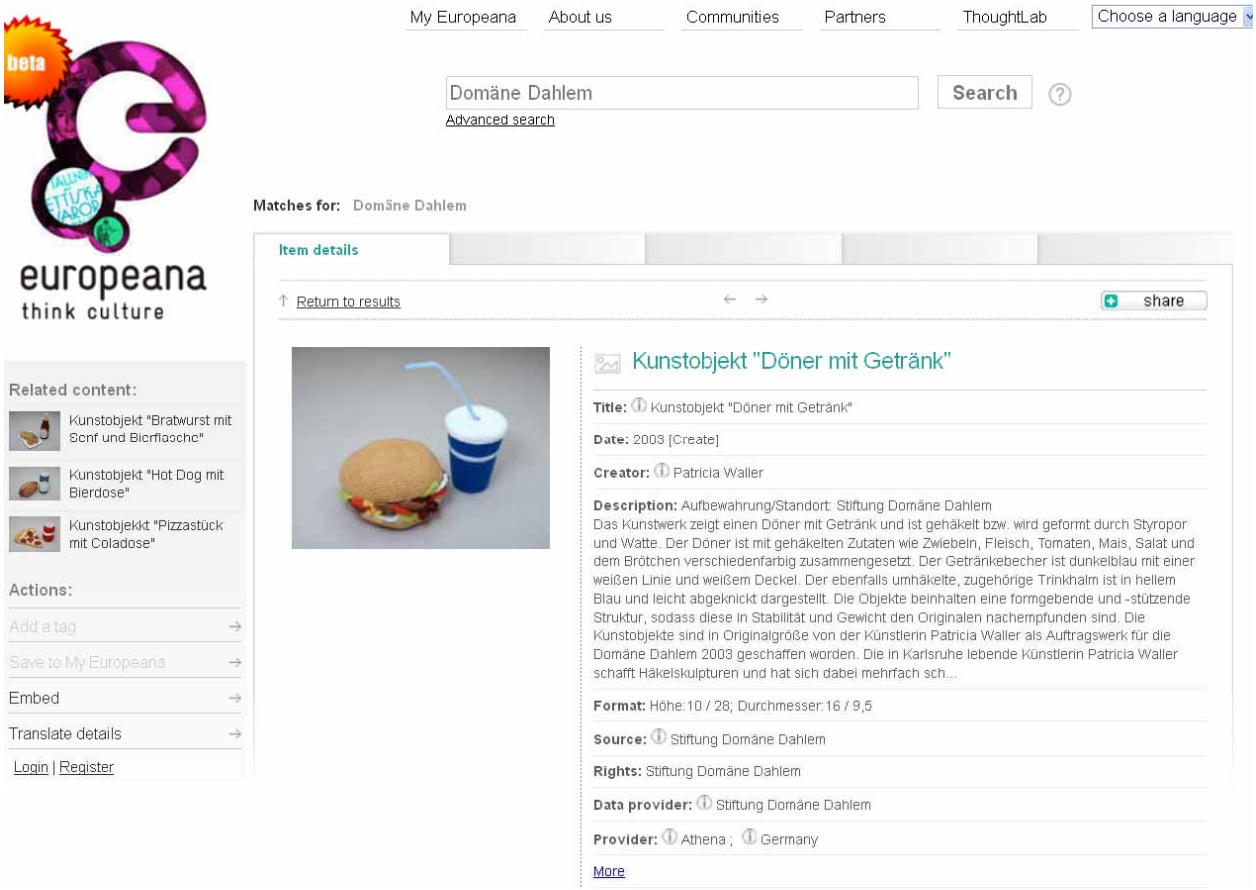

Few Charlemagne issued coins show the portrait of the emperor. This is the most beautiful coins under the portrait of Charlemagne. It was either marked the occasion of the coronation of the year from 800 to 812 or recognition of the empire of Charlemagne by Byzantium and shows us a life-temporal portrait of the emperor.

Otto-Lilienthal-Museum - Anklam
Wanddampfmaschine



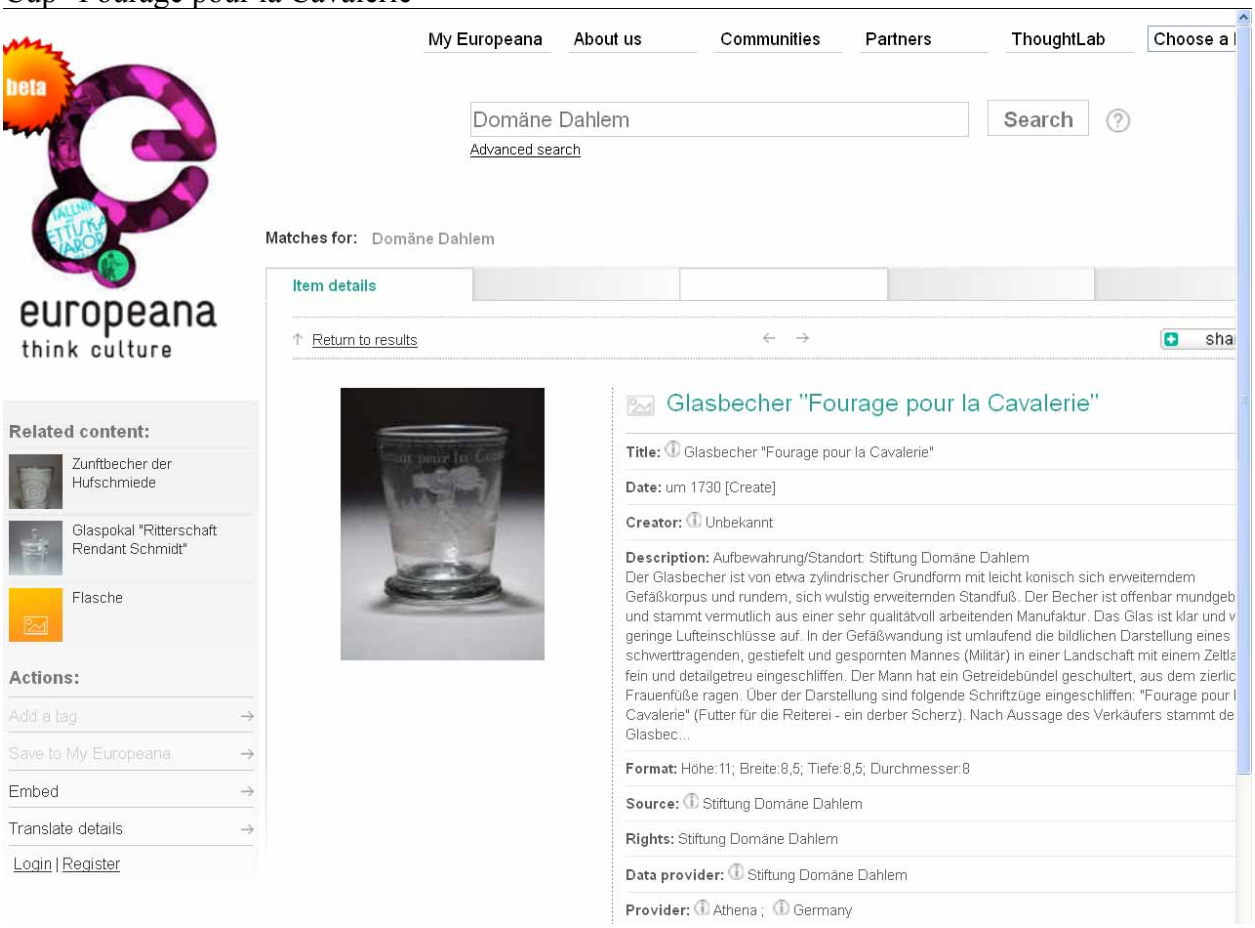

The acquisition of Lilienthal'schen Wanddampfmaschine (wall steam engine) from Australia was an internationally acclaimed sensation fund. It is the only remaining mechanical product of the factory "Otto Lilienthal" that was also just the first aircraft factory in the world.

Domäne Dahlem - Berlin
"Döner mit Getränk"



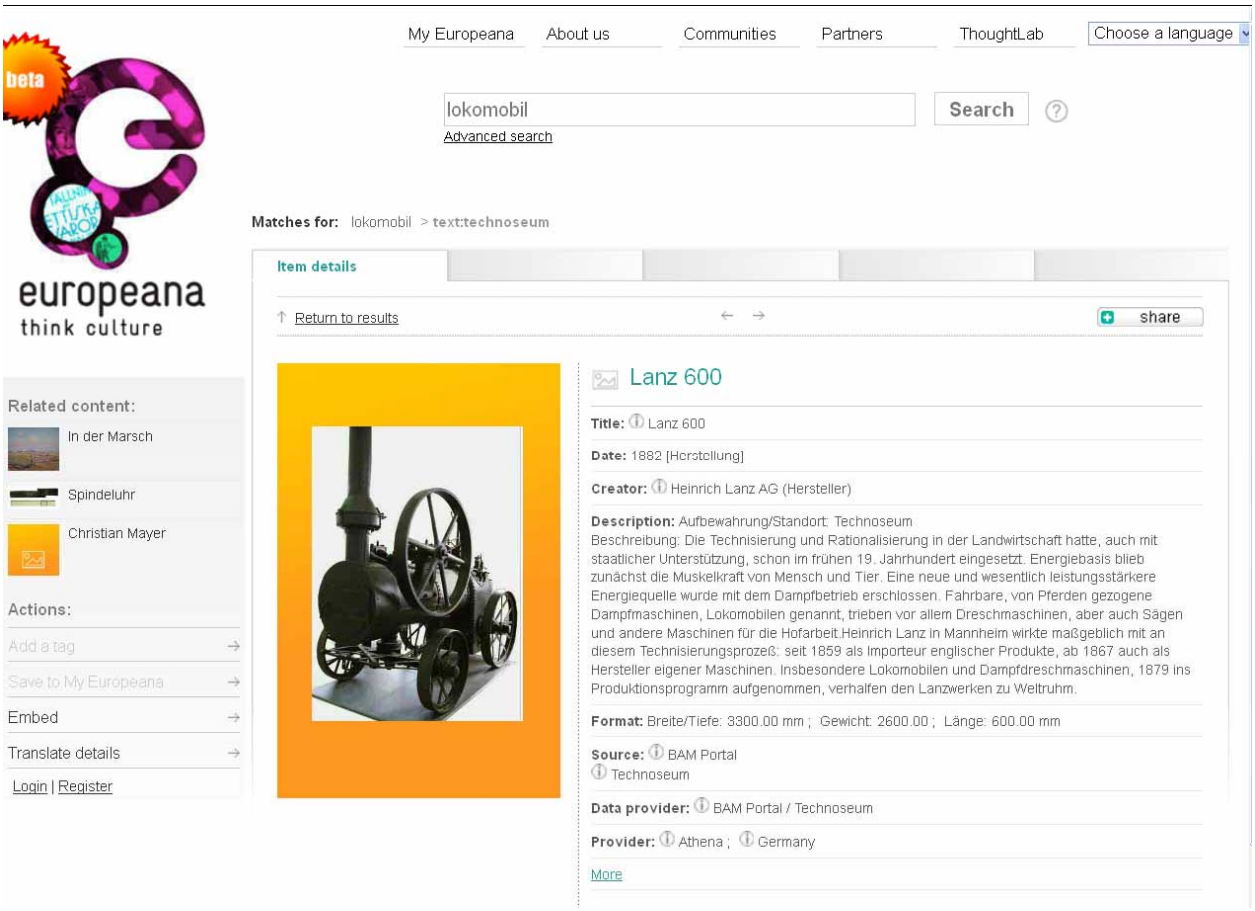

The artwork shows a kebab with a beverage and is crocheted and is formed by polystyrene and cotton wool. The art object created by the artist Patricia Waller within the "Eat Art"; it was produced as a commission for the exhibition "Imbissbuden" in 2003. Your crochet pieces offer a new perspective on well-known and distinguished by a humorous irony.

Cup "Fourage pour la Cavalerie"




The cup is from the Potsdam area, showing the connection of military history, agriculture, and supply troops. The ingrained word "forage pour la cavalry" (fodder for the cavalry) and the accompanying image of a military man dressed with a sheaf of wheat, revealed only on closer look the another meaning: not only the ears are carried away from the bundle also project female feet.

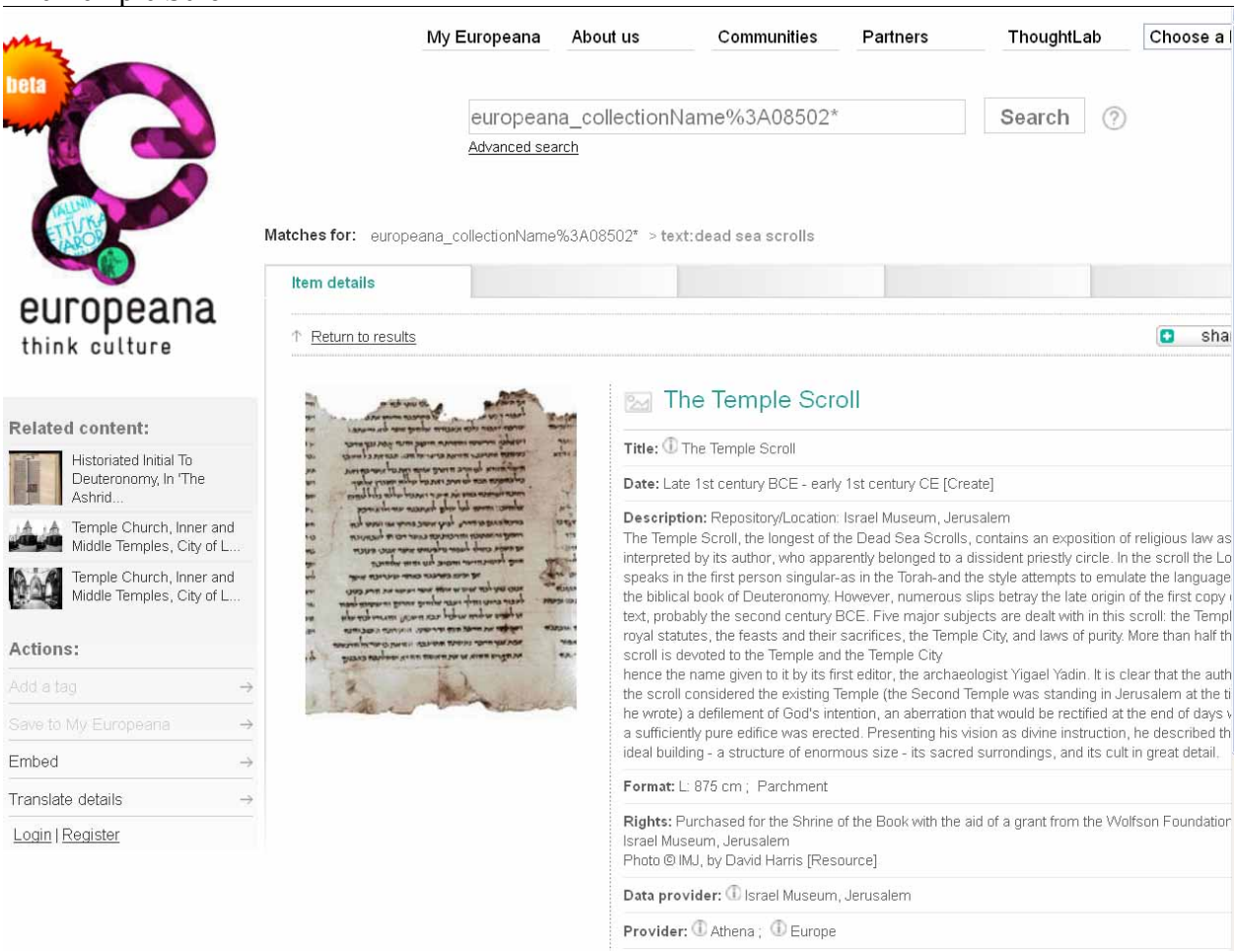
Technomuseum - Mannheim
Lokomobile Lanz




Built in Mannheim from 1879 by the Heinrich Lanz Company and sold all over the world, these portable steam engines were a milestone in the early transition process from traditional farming to industrial agriculture.

## ISRAEL

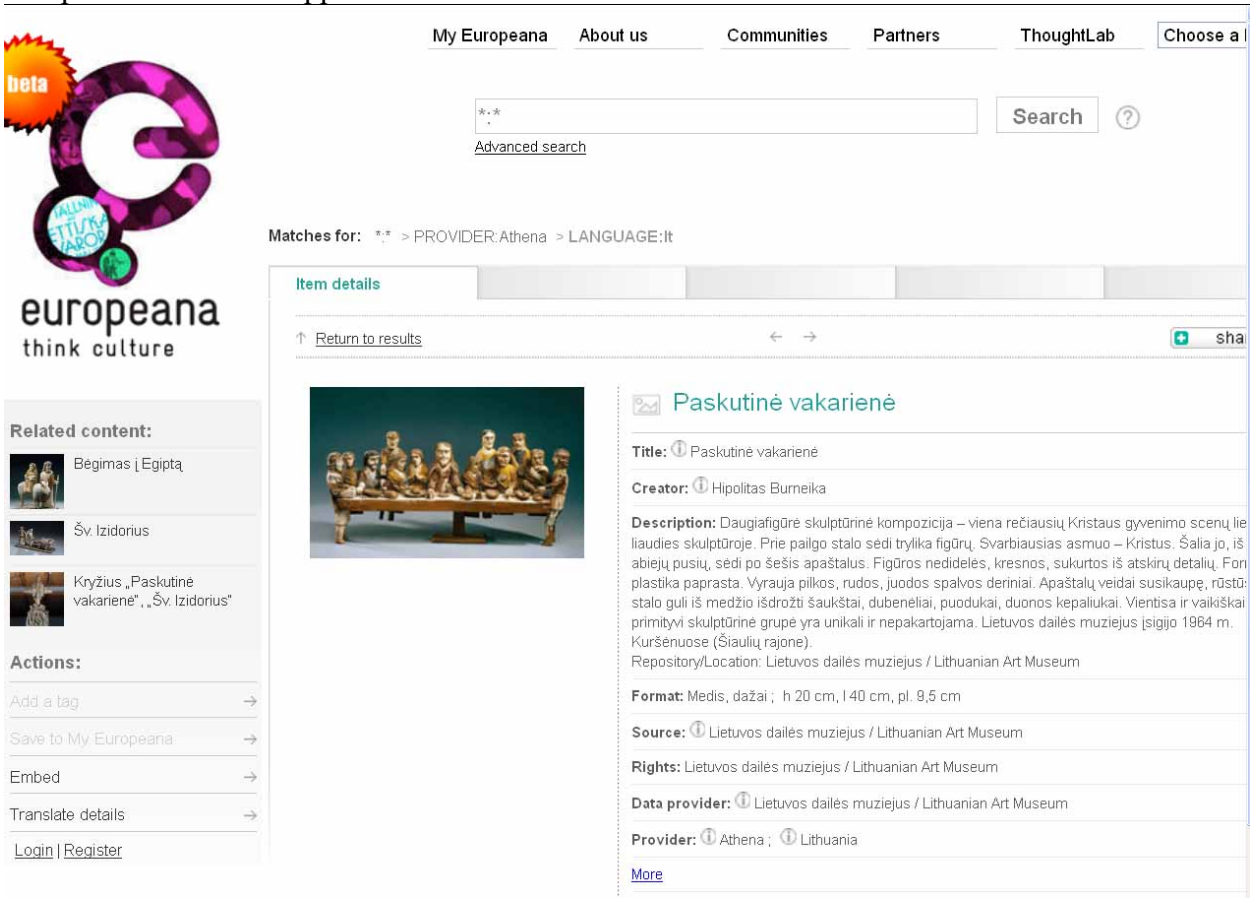
The Israel Museum, Jerusalem
The Temple Scroll




The Israel Museum's best known collection are the Dead Sea Scrolls, the manuscripts which are hosted in the Shrine of The Book. They include The Great Isaiah Scrolls, the oldest biblical manuscripts in the world — as well as several other rare early medieval biblical manuscripts. The Temple Scroll, which is the longest of the Dead Sea Scrolls, the Lord speaks in the first person singular, as He does in the Torah, and the style emulates the language of the biblical book of Deuteronomy.

## LITHUANIA

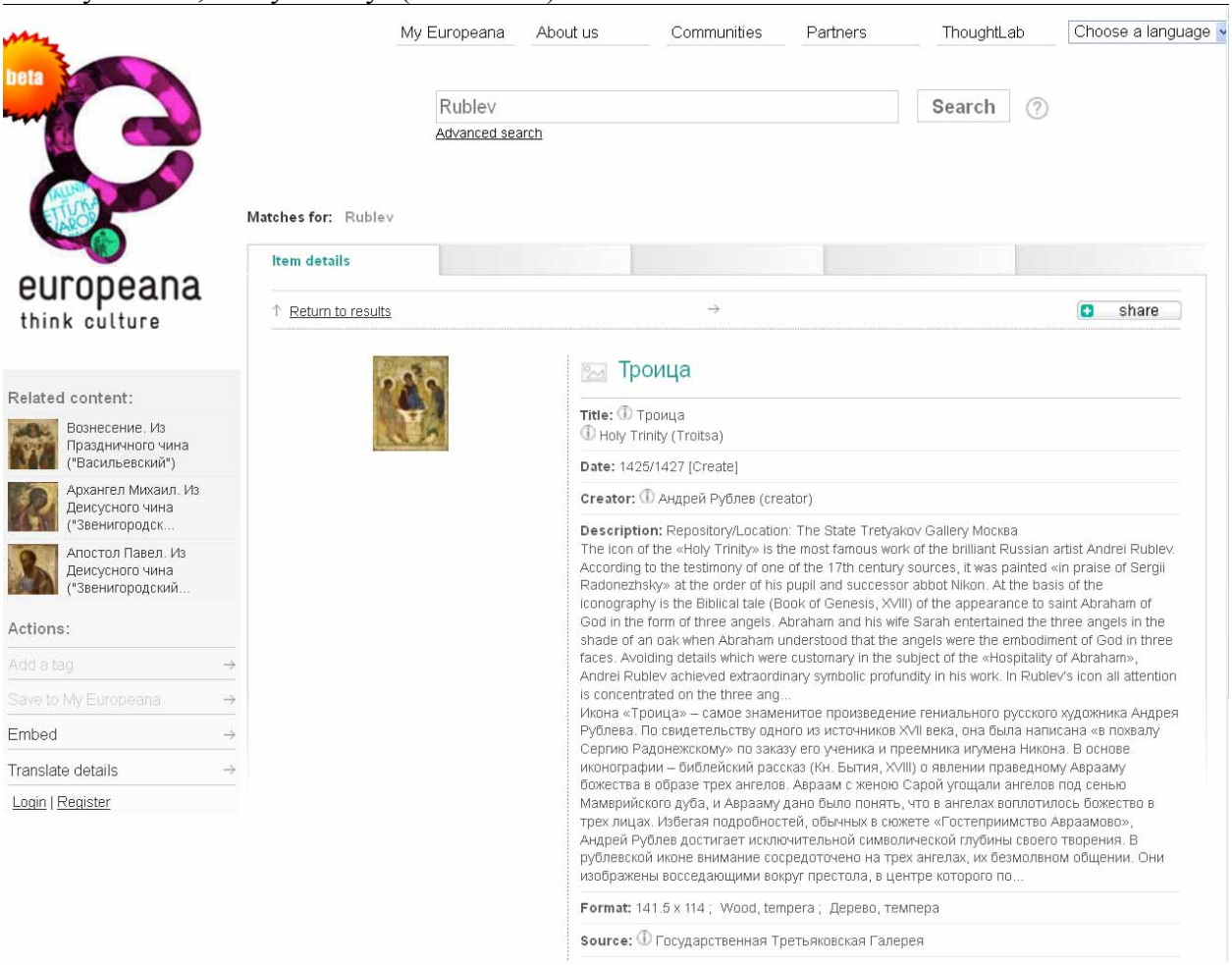
Lithuanian Art Museum
Sculpture “The Last Supper”




It is a piece of folk art sculpture made in the beginning of the 20th century by folk artist Hipolitas Burneika. The title is “The Last Supper”. Represented by the multi-figure sculptural composition, this particular scene of the life of Christ is one of the rarest examples of Lithuanian folk art. There are 13 figures sitting by the table – the Christ at the centre and apostles around him. The figures are small and stumpy, composed of separate parts. There are many small objects like spoons, plates and cups on the table. United, infantile and primitive this sculptural group is a unique work of art. It has been acquired by the Lithuanian Art Museum in 1964 from the village of Kuršėnai where it was made.

## RUSSIA

The State Tretykov Gallery
Andrey Rublev, 'Holy Trinity' (1425-1427)




The icon of the «Holy Trinity» is the most famous work of the brilliant Russian artist Andrei Rublev. According to the testimony of one of the 17th century sources, it was painted «in praise of Sergii Radonezhsky» at the order of his pupil and successor abbot Nikon. At the basis of the iconography is the Biblical tale (Book of Genesis, XVIII) of the appearance to saint Abraham of God in the form of three angels. Abraham and his wife Sarah entertained the three angels in the shade of an oak when Abraham understood that the angels were the embodiment of God in three faces. Avoiding details which were customary in the subject of the «Hospitality of Abraham», Andrei Rublev achieved extraordinary symbolic profundity in his work. In Rublev's icon all attention is concentrated on the three angels and their silent exchange. They are depicted as seated around an altar in the center of which there is a chalice of the Eucharist with the head of a sacrificial calf which symbolises the lamb of the New Testament, i.e., Christ. The left and centre angels bless the chalice. God the Father blesses God the Son for death on the cross in the name of love for people. God the Holy Spirit (the right angel) is present here to provide comfort, confirming the high logic of sacrificial, all-forgiving love. The content of the «Holy Trinity» is ambiguous. The monument is multi-faceted in its themes. Firstly, it embodies the idea of the triune Divinity. During the times of Sergii Radonezhsky and Andrei Rublev, the subject of the Trinity was understood as a symbol of spiritual unity, mutual love, the world and readiness to sacrifice oneself.

# Annex I: Ingestion plan

See the reserved area of the ATHENA website.