



europæana
think culture

When Semantics support Multilingual Access to Cultural Heritage

The Europeana Case

Valentine Charles and Juliane Stiller

SWIB 2014, Bonn, 2.12.2014



Our outline

1. Europeana
2. Multilinguality in digital libraries - challenges
3. Europeana Data Model – a framework for multilingual data
4. Semantic and multilingual enrichment



→ Europeana, the platform for
Europe's digital cultural heritage



Europeana

→ Aggregates metadata from the cultural heritage sector in Europe

- Libraries, museums, archives and audio-visual archives
- Metadata in 33 languages

→ Provides a portal for users to access data and objects

- <http://www.europeana.eu/> in 31 languages
- Metadata under Creative Commons Zero - public domain
- Previews and links to source

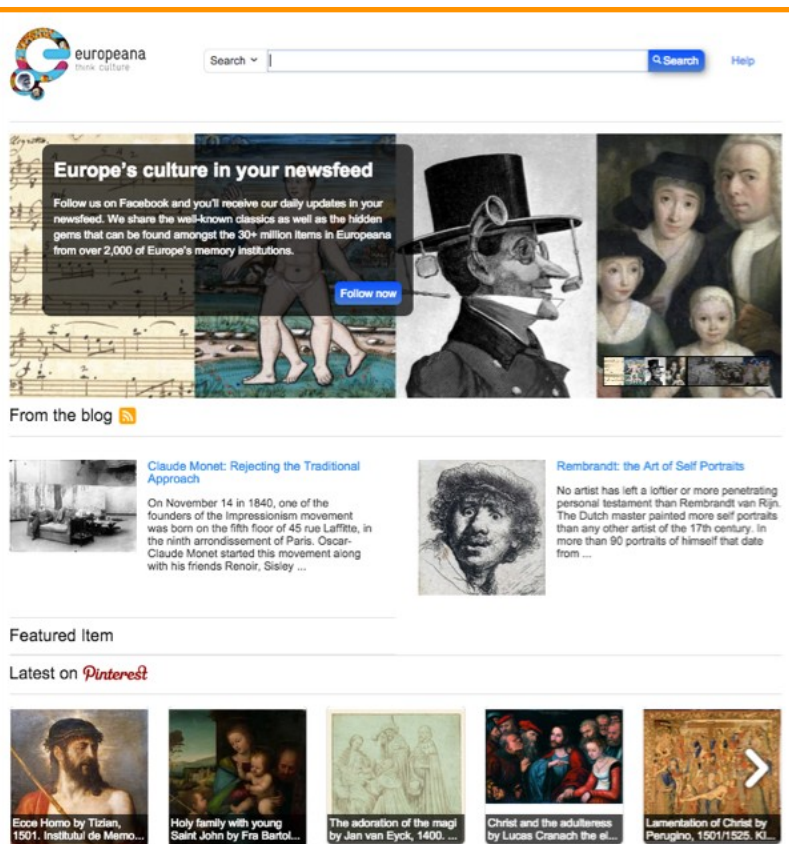
→ Data distributed via

- API <http://labs.europeana.eu/api/>
- Linked Data (currently being updated) <http://data.europeana.eu/>



Europeana.eu, Europe's cultural heritage portal

33M objects from 2,200 galleries, museums, archives and libraries



The screenshot shows the Europeana.eu homepage. At the top is the Europeana logo and a search bar. Below the search bar is a section titled "Europe's culture in your newsfeed" with a "Follow now" button. To the right of this is a "From the blog" section with two articles: "Claude Monet: Rejecting the Traditional Approach" and "Rembrandt: the Art of Self Portraits". Below these is a "Featured Item" section and a "Latest on Pinterest" section with five small image thumbnails.

Europeana
Share culture

Search

Europe's culture in your newsfeed

Follow us on Facebook and you'll receive our daily updates in your newsfeed. We share the well-known classics as well as the hidden gems that can be found amongst the 30+ million items in Europeana from over 2,000 of Europe's memory institutions.

Follow now

From the blog

Claude Monet: Rejecting the Traditional Approach

On November 14 in 1840, one of the founders of the Impressionism movement was born on the fifth floor of 45 rue Laffitte, in the ninth arrondissement of Paris. Oscar-Claude Monet started this movement along with his friends Renoir, Sisley ...

Rembrandt: the Art of Self Portraits

No artist has left a loftier or more penetrating personal testament than Rembrandt van Rijn. The Dutch master painted more self portraits than any other artist of the 17th century. In more than 90 portraits of himself that date from ...

Featured Item

Latest on Pinterest

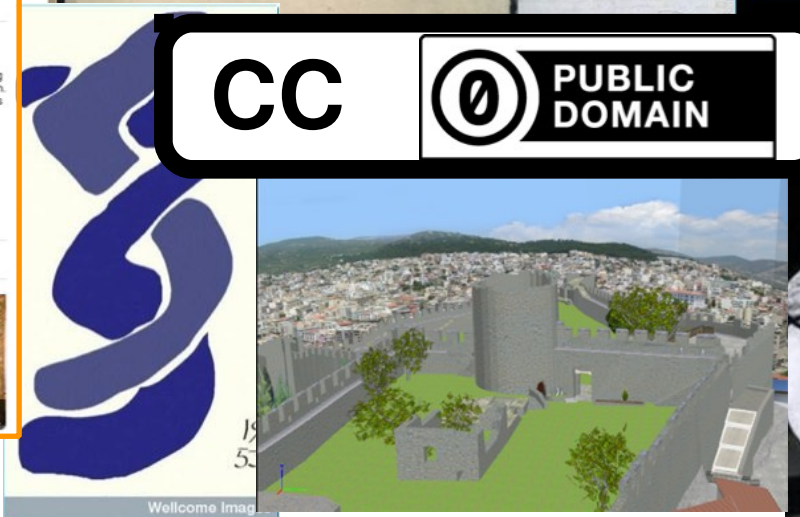
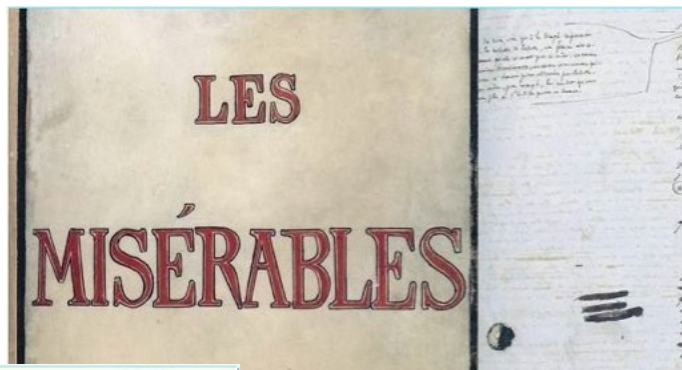
Ecco Homo by Titian, 1501. Institutul de Memo...

Holy family with young Saint John by Fra Bartol...

The adoration of the magi by Jan van Eyck, 1400...

Christ and the adulteress by Lucas Cranach the el...

Lamentation of Christ by Perugino, 1501/1525. KI...



Challenges

➡ Multilinguality issues

- Provide access to multilingual resources
- Allow the search for items in various languages
- Make sure users can understand the descriptions of these items



→ Multilinguality in digital libraries - challenges



Dimensions of multilinguality

- Interface and portal display
- Search
 - Translation of query
 - Translation of documents
- Representation and refinement of search results
 - User needs to be able to determine relevance of documents
- Browsing



Portal display

It looks like your browser is set to German; Would you also like to use Europeana in German? [yes](#) / [no](#)
Offensichtlich arbeitet ihr Webbrowser auf deutsch. Nutzen Sie auch Europeana lieber auf deutsch?? [ja](#) / [nein](#)

Home My Europeana



Search ▾

Europe's culture in your newsfeed

Follow us on Facebook and you'll receive our daily updates in your newsfeed. We share the well-known classics as well as the hidden gems that can be found amongst the 30+ million items in Europeana from over 2,000 of Europe's memory institutions.

Choose a language ▾

Choose a language

English

Basque

Български

Català

Če

Dansk

Deutsch

Ελληνικά

Español

Eesti

Français

- Which language will be displayed to the (first) user?
- Will a cookie be set?
- What will be translated?
- Which language dimensions does the drop-down menu impact?



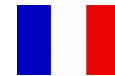
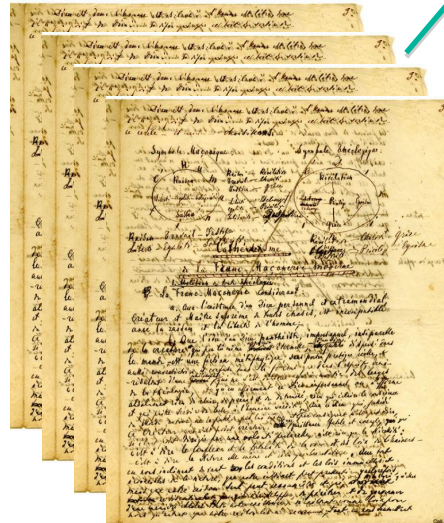
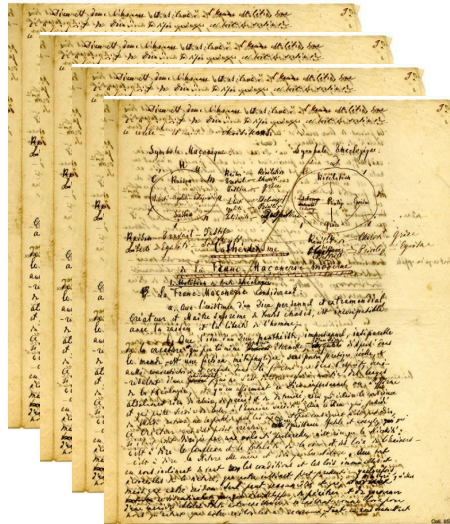
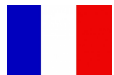
Cross-lingual search

Mona Lisa AND La Joconde

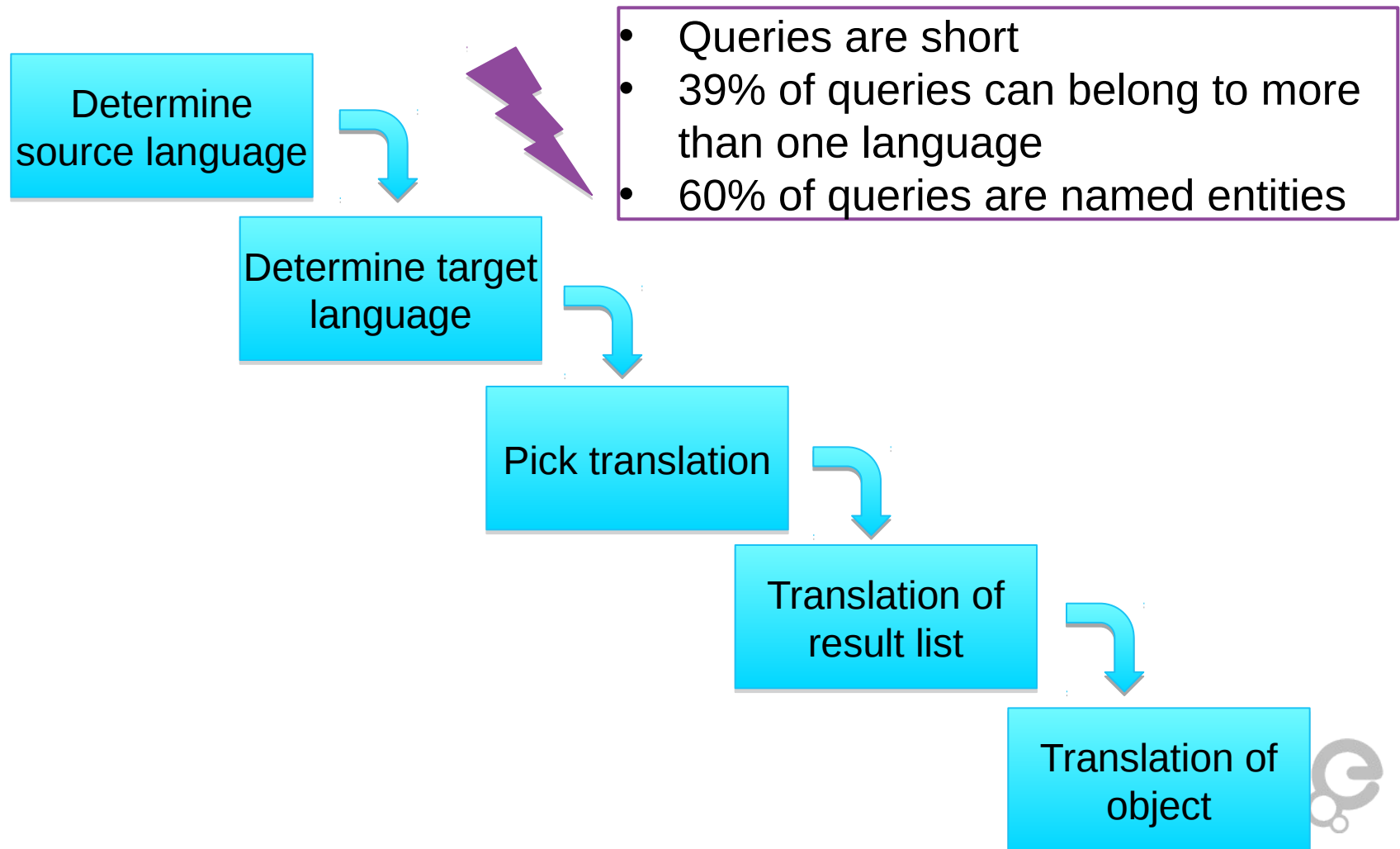
Search



External Dataset



Cross-lingual search



→ Europeana Data Model- a framework for multilingual data



Create new data framework

→ Europeana Data Model (EDM)

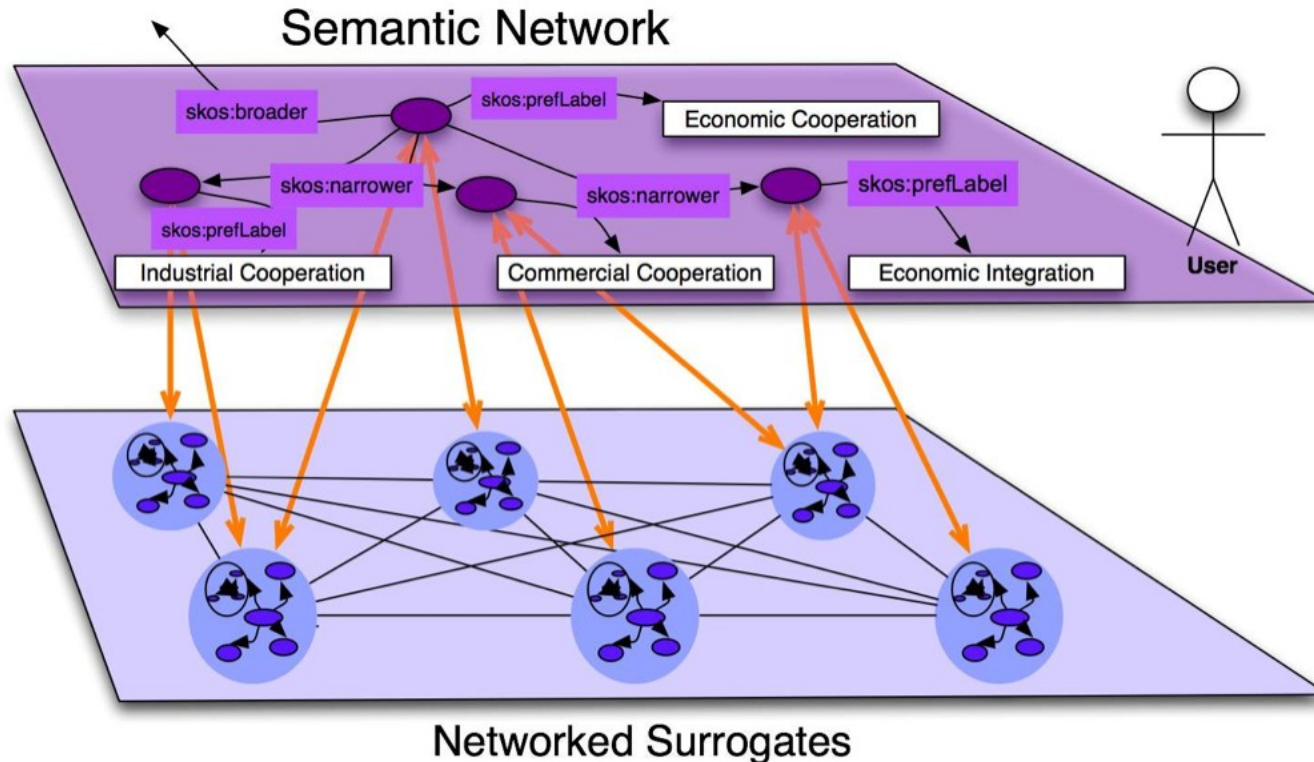
- Re-uses several existing Semantic Web-based models: Dublin Core, OAI-ORE, SKOS, CIDOC-CRM...

→ More granular metadata

- Links e.g. between objects and context entities (persons, places)
- Multilingual & semantic linked data for contextual resources (e.g. Concepts)



Rely on knowledge organisation systems



→ Create a “semantic layer” on top of cultural heritage objects

- Include multilingual “value vocabularies”
- From Europeana’s providers or from third-party data sources

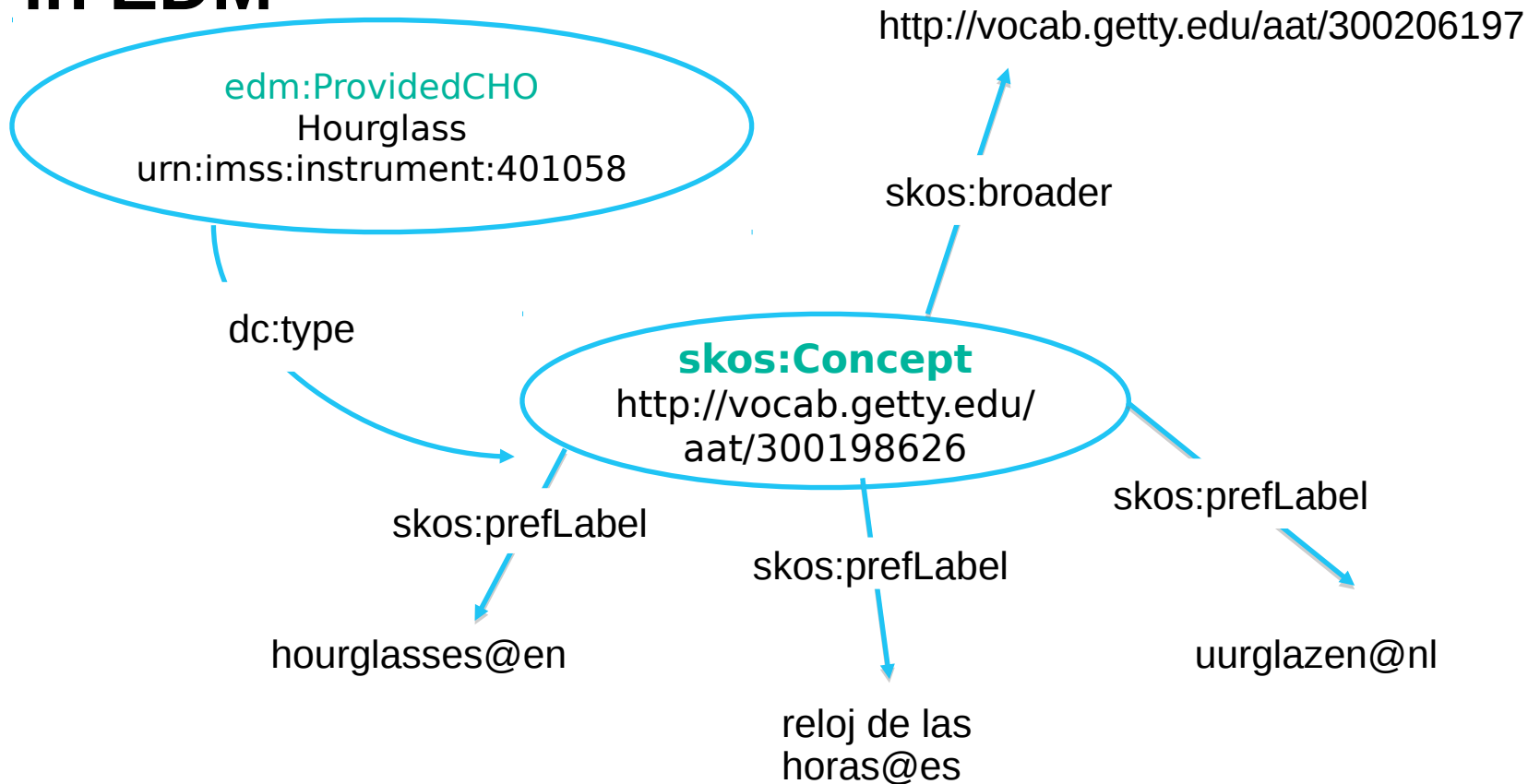


Encourage providers to contribute their own vocabularies

- Benefit from data links made at data providers' level
- Ingestion of vocabularies is made possible if the vocabularies used the data structures EDM expects
 - For instance SKOS for concepts



An example the integration of AAT URIs in EDM





© Free access - no re-use

View item at
[Museo Galileo - Istituto e Museo di Storia della Scienza](#)



Share



Cite on Wikipedia



Translate details

Select language

Powered by Microsoft® Translator

Clessidra

Title:

Hourglass

Description:

Clessidra a polvere racchiusa in una incastellatura esagonale di ferro con sei piccole colonne. Le capsule di vetro sono unite al centro da cerchi di carta. Il tempo trascorso veniva misurato dalla quantità di polvere passata dall'ampolla superiore a quella inferiore.;

Powder hourglass housed in a hexagonal iron frame with six small pillars. The glass compartments are joined at the center by paper disks. The time elapsed was measured by the quantity of powder falling from the upper compartment to the lower one.

Date of creation:

sec. XVII; 1601; 1700; 17th cent.

Type:

[Clessidra](#); [hourglasses](#); [hourglass](#); [hour-glasses](#)

Format:

altezza 130 mm; height 130 mm; ferro, vetro; [glass \(material\)](#); [iron \(metal\)](#); [ferrous metal](#); iron, glass

Identifier:

urn:imss:instrument:401058

Relation:

<http://catalogue.museogalileo.it/indepth/HourglassesWaterClocks.html>

Language:

eng ita

Publisher:

Museo Galileo - Istituto e Museo di Storia della Scienza

Data provider:

[Museo Galileo - Istituto e Museo di Storia della Scienza](#)

Provider:

[Museo Galileo - Istituto e Museo di Storia della Scienza](#)

Providing country:

Italy

Auto-generated tags ▾

What ▾

Concept Term:

<http://vocab.getty.edu/aat/300198626>

Concept Label:

[hourglasses] (en); [reloj de las horas] (es); [uurglazen] (nl)

Concept Broader Label:

Broader concept: <http://vocab.getty.edu/aat/300041573> ↗

Narrower concept: [hourglasses](#)

Concept Term:

<http://vocab.getty.edu/aat/300206197>

Concept Label:

[sabliers] (fr); [sandglasses] (en); [reloj de arena] (es); [zandlopers] (nl)

Concept Broader Label:

<http://vocab.getty.edu/aat/300041573>

Concept Term:

<http://vocab.getty.edu/aat/300010797>



→ Automatic enrichments

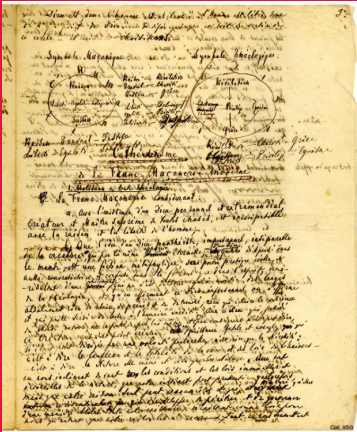


Enrichments in information retrieval

Mona Lisa AND La Joconde

Search

Object

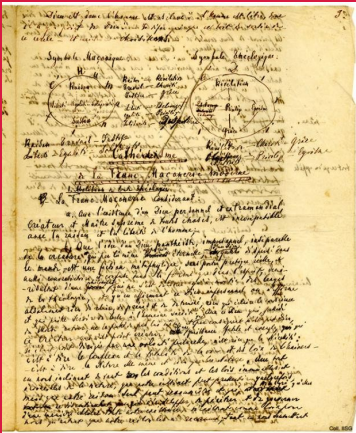


Goal: reaching higher visibility of documents within the document space



Enrichments in the linked data space

Object



External Dataset
and Vocabulary

Goal: contextualization which goes beyond the scope of a particular platform



Automatic enrichment process in Europeana

1. Analysis

- Metadata fields in resource descriptions
- Selection of potential rules to match

2. Linking

- Matching the values of the metadata fields to values of the contextual resources
- Adding contextual links

3. Augmentation

- Selecting the values from the contextual resource
- Values go into the search index



Enrichment types and vocabularies

Enrichment Type	Target vocabulary	Source metadata fields	Number of enriched objects
Places	GeoNames	dcterms:spatial, dc:coverage	7 mio
Concepts	GEMET, DBpedia,	dc:subject, dc:type	9,2 mio
Agents	DBpedia	dc:creator, dc:contributor	144,000
Time	Semium Time	dc:date, dc:coverage, dcterms:temporal, edm:year	10,2 mio

Europeana Enrichment - Example

[Return to search results](#)

[< Previous](#) [Next >](#)



© Free access - no re-use

View item at
[National Library of the Netherlands - Koninklijke Bibliotheek](#)



Share



Cite



Translate

Select language

Powered by

De koppelaarster

Description: Litt. : Christopher Braider: Refiguring the real. Picture and modernity in word and image. 1400-1700. Princeton, 1993

Creator: Johannes Vermeer ;
http://dbpedia.org/resource/Johannes_Vermeer

Contributor: Johannes Vermeer

Date of creation: 16XX

Type: StillImage

Format: schilderij ; image/jpeg

Subject: koppelen

Identifier:
<http://resolver.kb.nl/resolve?urn=urn:gvn:NCRD01:107484099>

Search also for:

Title

[De koppelaarster](#) (7)

Who

[Johannes Vermeer](#) (8)
http://dbpedia.org/resource/Johannes_Vermeer (3)
[jan vermeer](#) (5)
[扬·弗美尔](#) (3)
[jan vermeer](#) (5)
[johannes vermeer](#) (8)
[jan vermeer](#) (5)
[johannes vermeer](#) (8)
[johannes vermeer](#) (8)
[johannes vermeer](#) (8)
[вермеер, ян](#) (3)
[johannes vermeer](#) (8)
[johannes vermeer](#) (8)

What

[Auto-generated tags](#) ▾

[Who](#) ▾

Agent Term: http://dbpedia.org/resource/Johannes_Vermeer

Agent Label: [jan vermeer] (de) ; [扬·弗美尔] (zh) ; [jan vermeer] (it) ; [johannes vermeer] (pt) ; [jan vermeer] (pl) ; [johannes vermeer] (sv) ; [johannes vermeer] (fr) ; [johannes vermeer] (en) ; [вермеер, ян] (ru) ; [johannes vermeer] (es) ; [johannes vermeer] (nl)

Providing country: Netherlands

[Auto-generated tags](#) ▾

[Who](#) ▸

Quality of enrichments

- Olensky et al. (2012) analyzed 200 enrichments of Europeana
-> found enrichment flaws and problems
- Incorrect enrichments lead to
 - Devaluation of curated metadata
 - Loss of trust from providers
 - Propagation of errors to different languages
 - Irrelevant search results
 - Bad user experiences



Better understanding of impact of enrichments needed



To conclude

- Continue to focus on cross-domain multilingual vocabulary alignment and publish the results as Linked Data
 - More pivot vocabularies such as AGROVOC, STW Thesaurus for Economics integrated in Europeana
- More domain-specific and targeted vocabularies for enrichment
- Multilingual interactions
- Better understanding of impact of multilingual strategies on Search and Browse and User Interactions





europeana
think culture

Thank you

Valentine Charles & Juliane Stiller

valentine.charles@europeana.eu, juliane.stiller@ibi.hu-berlin.de

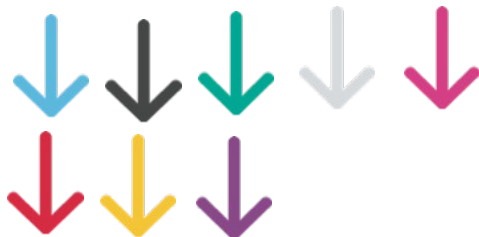
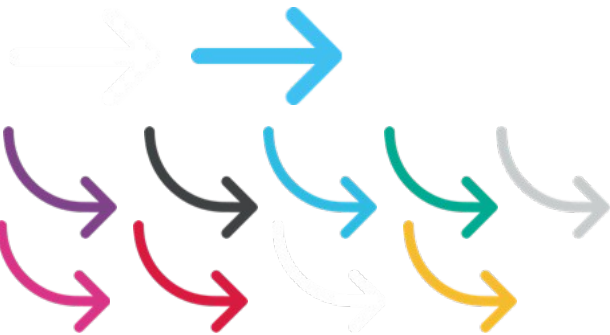


Toolbox

1 2 3 4 5 6 7 8 9 0

1 2 3 4 5 6 7 8 9 0

1 2 3 4 5 6 7 8 9 0



Replace text and
adjust size

Replace text and
adjust size

Replace text and
adjust size

Replace text and
adjust size

Replace text and
adjust size

Replace text and
adjust size

Aggregate

Distribute

Engage

Facilitate

