ATHENA PLUS
Access to cultural heritage networks for Europeana

European Commission

cip competitiveness and innovation framework programme 2007–2013

AthenaPlus

WP 4

Terminologies and semantic enrichment

How to manage multilingual terminologies for a better access to the network

Marie-Véronique Leroi & Eva Coudyzer

AthenaPlus workshop/TEI Conference

Outline

1. Terminology?

2. Semantic Web

3. AthenaPlus WP4

6. Terminology Management Platform (TMP)

Terminology?

Who?

- Galleries, libraries, archives and museums (GLAM) use monolingual or multilingual terminologies.

Definitions

- Aitchison et al. (2000): “a vocabulary of controlled indexing language, formally organized so that a priori relationships between concepts are made explicit”.

- Hodge (2000): “Knowledge Organization Systems are used to organize materials for the purpose of retrieval and to manage a collection. A KOS serves as a bridge between the user’s information need and the material in the collection. With it, the user should be able to identify an object of interest without prior knowledge of its existence. Whether through browsing or direct searching, whether through themes on a web page or a site search engine, the KOS guides the user through a discovery process”.

Terminology?

You need to use a list of terms explicitly enumerated → SIMPLE LIST OF TERMS

\+ *definitions* =

You need to use a list of terms with definitions or explanations → GLOSSARY

\+ *hierarchy* =

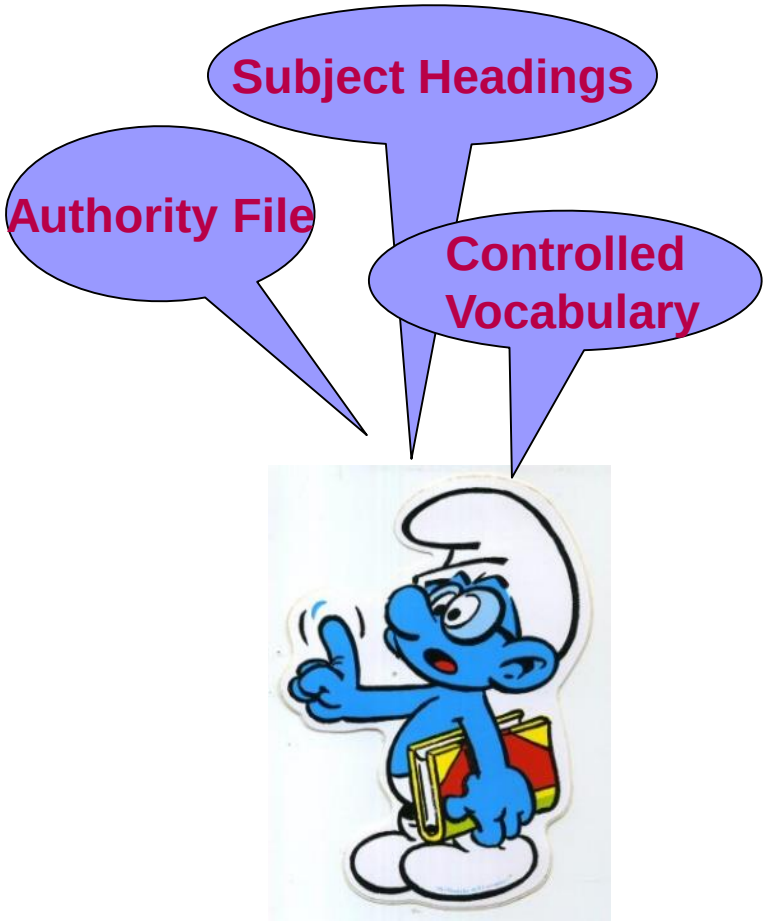
You need to use a list of terms in a hierarchical structure → CLASSIFICATION / TAXONOMY

\+ *semantic relations* =

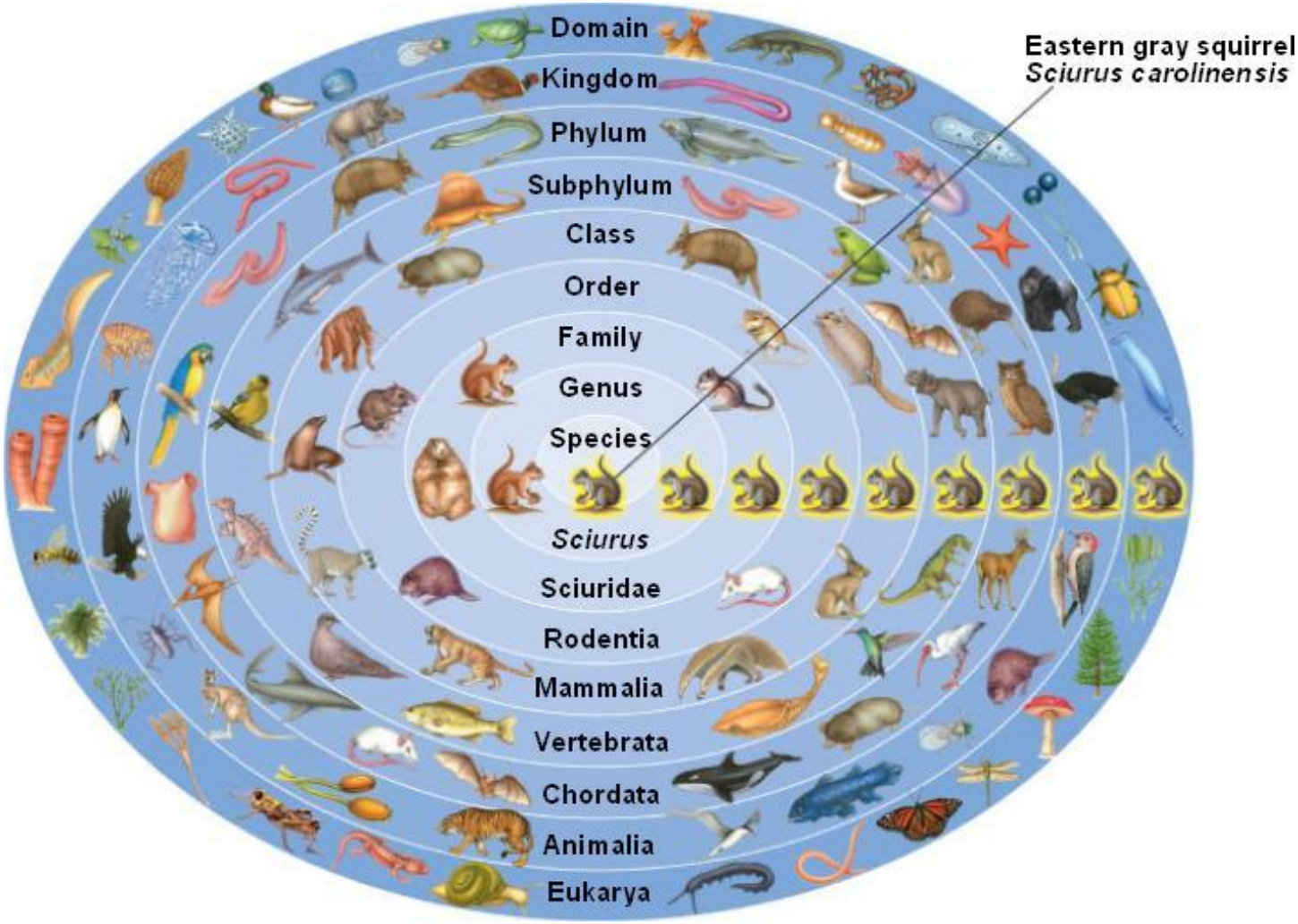
You need to use a network collection of controlled vocabulary terms → THESAURUS

\+ *concepts* =

You need to use a formal representation of a set of concepts → ONTOLOGY



Some examples :
Linnean taxonomy of organisms



Some examples :
Getty thesaurus of Geographic Names

ID: 7008038 **Record Type: administrative**

Paris (inhabited place)

Coordinates:
Lat: 48 52 00 N *degrees minutes* Lat: 48.8667 *decimal degrees*
Long: 002 20 00 E *degrees minutes* Long: 2.3333 *decimal degrees*

Note: Paris is the city and capital of France, located in the north-central part of the country. It was founded more than 2,000 years ago on an island in the Seine River. The City of Paris itself covers an area of 41 square miles (105 square kilometers). It occupies a central position in the productive agricultural region known as the Paris Basin. It is the nation's most important center of commerce and culture. It was captured and fortified by the Romans in 52 BCE and ruled by various other groups such as the Clovis, Northmen, English, and Germans.

Names:

Paris (**preferred**,C,V,English-P,U,N)
(French-P,U,N)
Parigi (C,O,Italian-P,U,N)
París (C,O,Spanish-P,U,N)
Parisii (H,V)
Lutetia (H,V) ..... ancient name
Lutetia Parisiorum (H,V,Latin,U,N) ..... ancient
Lutetia Parisii (H,V)
Parisius (H,V)
parisien (C,V,French,U,A)
parisienne (C,V,French,U,A)
Parisian (C,O)

Hierarchical Position:

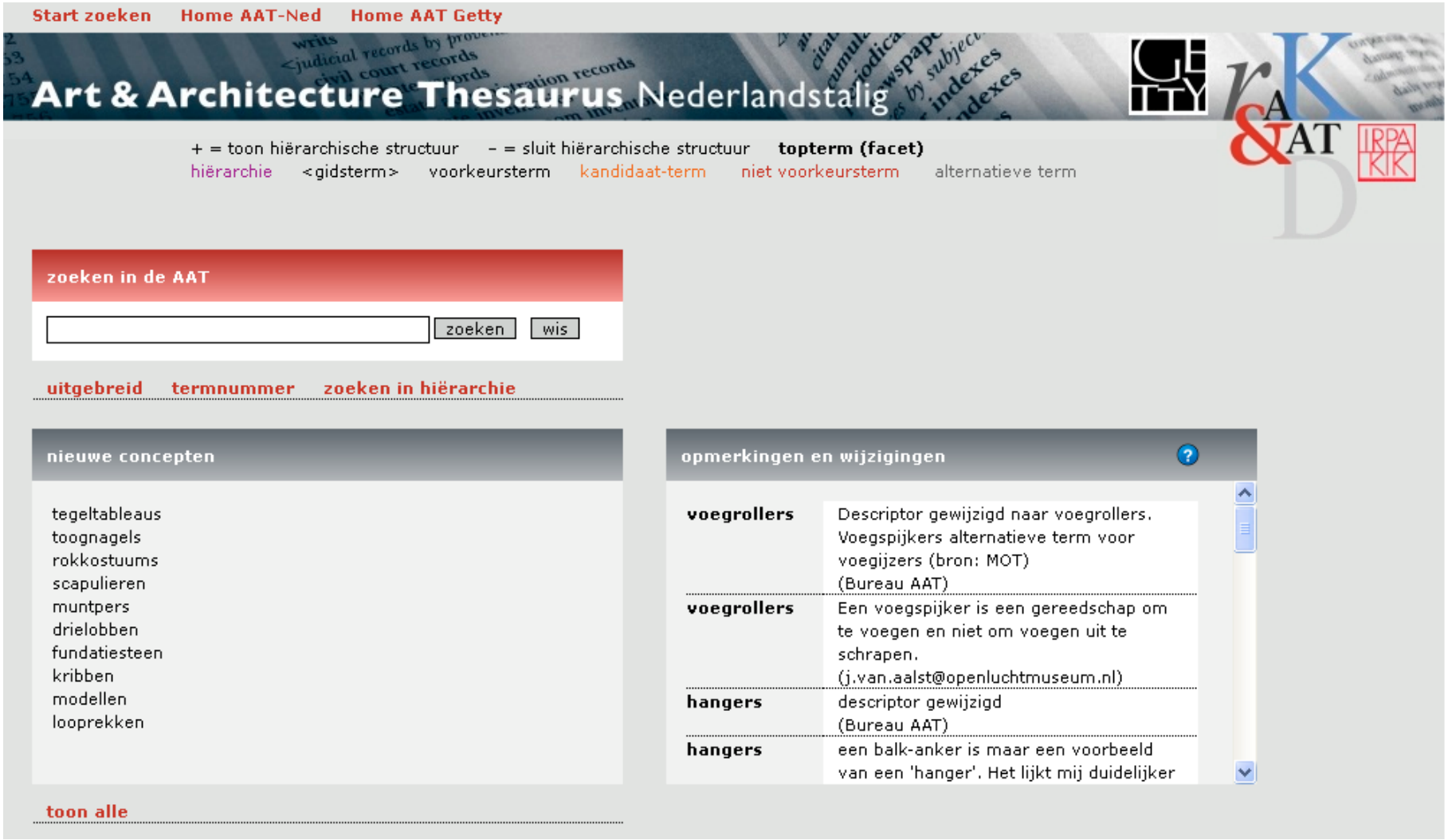
World (facet)
.... Europe (continent) (P)
..... France (nation) (P)
..... Île-de-France (region) (P)
..... Ville de Paris, Département de (department) (P)
..... Paris (inhabited place) (P)

Additional Parents:

World (facet)
.... Europe (continent) (P)
..... France (nation) (P)
..... Île-de-France (historical region) (P)
..... Paris (inhabited place)

World (facet)
.... Europe (continent) (P)
..... Gaul (historical region) (P)
..... Gallia Lugdunensis (province) (P,H)
..... Paris (inhabited place) (P,H)

Some examples :
Getty Art & Architecture Thesaurus (NL)



Why do GLAM-sectors use thesauri?

Use as :

- Indexing tool
- Search tool / query formulation support feature
- Interactive term suggestion
- Browse and navigation tool
- To facilitate combination of multiple databases or unified access to multiple databases

Role :

- Thesauri have been a primary dimension of the research and development interests of experts in information retrieval, online searching, user interface design, knowledge organization in general, indexing and abstracting, cataloguing and classification, and information search behavior studies (Ali Shiri)

The Semantic Web

What?

- Introduced in 2001 by Tim Berners-Lee
- Collaborative effort led by W3C

Principle? share and reuse data on the Web (semantic interoperability)

Why? better search results, irrespective of language

How? By automatically linking “separate” data on the Web (linked data); by inter-thesauri mapping

Tools? Allow machines to understand the meaning, or “semantics”, of information on the Web

Thesauri can serve as a basis for developing domain-specific ontologies

Semantic Web and Linked Data

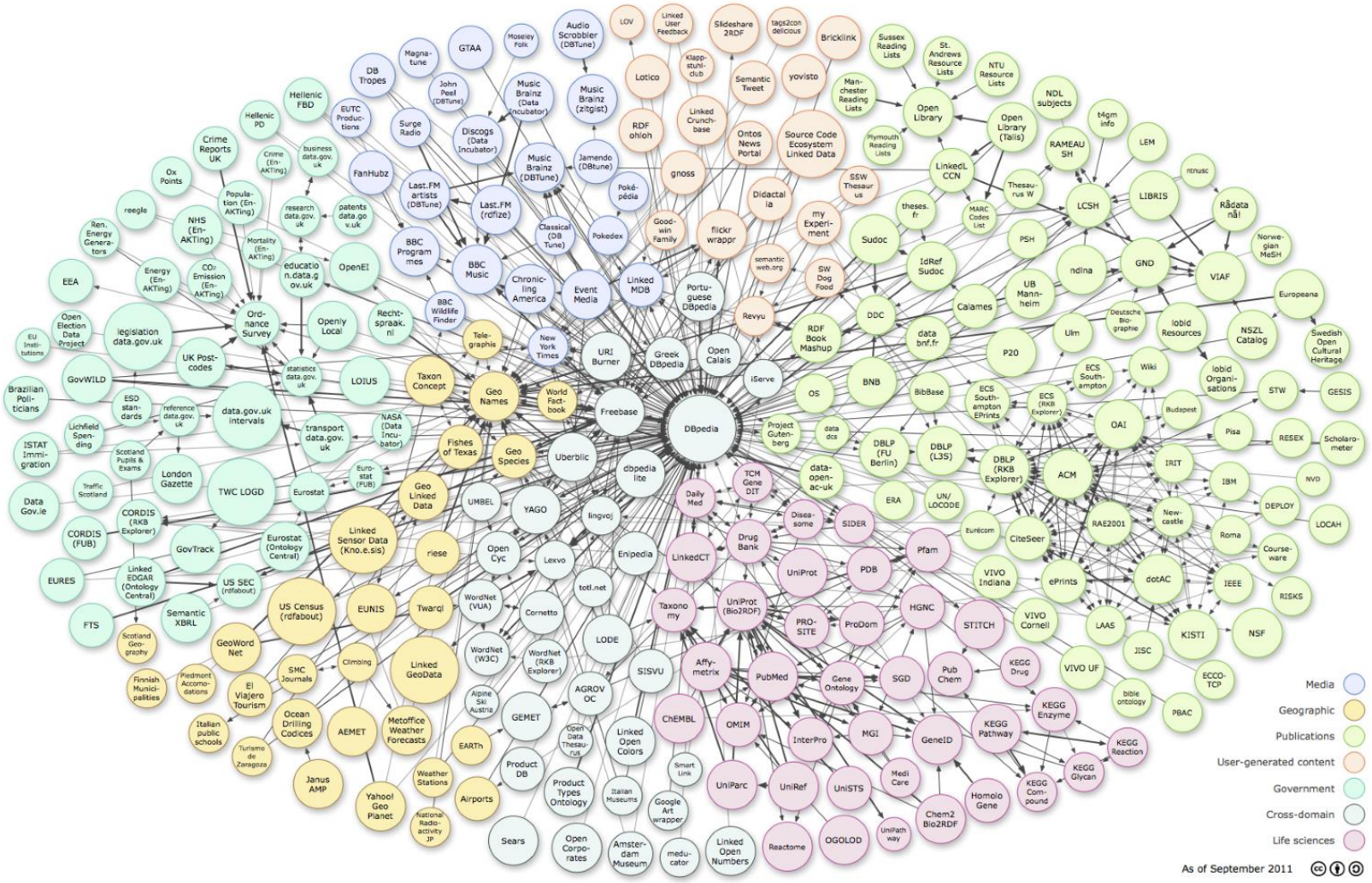
Semantic Web

"the Web of data with meaning in the sense that a computer program can learn enough about what the data means to process it" (Tim Berners-Lee)

Linked Data

"The Semantic Web isn't just about putting data on the web. It is about making links, so that a person or machine can explore the web of data. With linked data, when you have some of it, you can find other, related, data." (Tim Berners-Lee)

Semantic Web and Linked Data



Why joining the Semantic Web?

- Keep providing the same quality of service as a public institution and improve it

- Give access to structured and normalised cultural data

- Be part of the Linked Open Data Cloud

- Enhance the semantic interoperability at national and international level

- Enhance the visibility of your collections and professionnal knowledge

- Help the user in getting optimised results to specific requests

- Be a trusted source of cultural data

European CIP-Best practice network: AthenaPlus

- 2013-2015

- WP4 dedicated to terminology and semantic enrichment

- WP Leaders : KMKG and MCC
- Technical partner : University of Savoie
- WP Working group

- Legacy

- Athena, eContentPlus project : 2008-2011;
- Linked Heritage, CIP Best practice network : 2011-2013
- Activity on the working groups of the Michael Culture Association

General objectives

- Develop and provide a fully functionnal Terminology Management Platform (TMP)
- Organise training and awareness on terminology and the use of TMP
- Provide the semantic enrichment of metadata and terminologies

Specific objectives

- Set up a registry to serve as an archive of reference terminologies stored in an Europeana interoperable format (SKOS RDF)
- Provide a web service for online sharing of terminologies and semantic mappings between concepts of different terminologies in order to publish the terminologies as part of the Linked Data Web
- Provide education and training to Content Providers on best practice multilingual terminology management, SKOSification, semantic mapping of concepts and Linked Data, publication of terminologies using the dedicated AthenaPlus Terminology Management Platform

Terminology resources

From :

- AthenaPlus partners
- Linked Heritage partners
- project contributors, …

- 20 European countries
- 18 different languages
- GLAM-sector

Important criteria :

- Controlled vocabulary?
- Multilingual?
- Persistent identifiers?
- Free of rights?

Terminology questionnaire

- Organization
- Terminology type
- Management
- Standards
- Export formats
- Target groups
- Accessibility
- Rights
- Semantic enrichment

33 partners

62 vocabularies

Questionnaire results

- Terminologies are controlled
- Classification, thesaurus, authority files
- Some combination of more than one type
- Some: person/organization names (FOAF is better model)
- Terminologies are well managed
- Notes (e.g. scope notes) are rare

- Not many multilingual vocabularies
- Many different domains

Questionnaire results

Metadata:

- Object names
- Material
- Techniques
- Geographic locations
- Authors
- Organizations
- Iconography
- Style/period
- Collection type

Results

- **In-house:**
  - Thesaurus of Museum of Arts and Crafts (Zagreb)
  - Object name thesaurus of Royal Museums of Art and History (Brussels)
  - Theatre Institute Library Classification of Specialized Literature (Prague)
  - …
- **Adopted/bought:**
  - GND
  - DDC
  - AAT
  - PICO 4.3
  - …

Questionnaire results

Rights:

- Terminologies under CC-licence
- Terminologies free of rights but no specific licence
- Terminologies not free of rights
- Rights have not been discussed internally

Athena/Linked Heritage Booklet

cip

LINKED HERITAGE

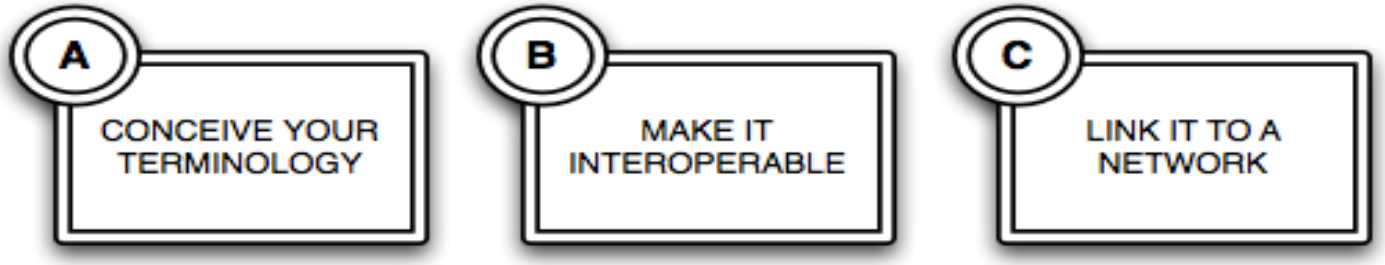
europeana
project group member

YOUR TERMINOLOGY
AS A PART OF THE SEMANTIC WEB
RECOMMENDATIONS
FOR DESIGN AND MANAGEMENT

MINISTERO PER I BENI E LE ATTIVITÀ CULTURALI

Michael Culture Association

Athena : Final recommendations



A : What is the best way to create a terminology?

Create or adapt an existing terminology

B : What format do I use to make my terminology interoperable?

Technical implementation of the created or adapted terminology (in the perspective of the Semantic Web)

C : How can I link my terminology to other terminologies?

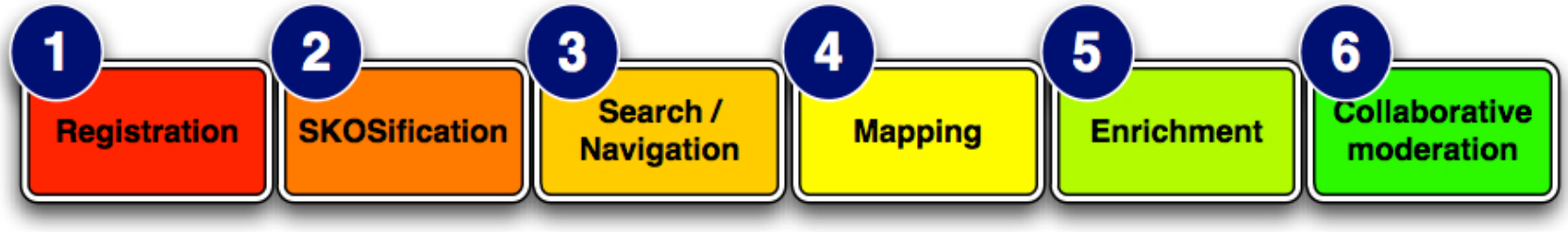
Linking of the SKOSified terminology with other resources (in the perspective of Linked Open Data)

Athena benchmark

How to connect and build bridges in-between multilingual vocabularies?

Functional needs identified by Athena WP4: Benchmark

- Registration of a terminology in a repository
- SKOSification of a terminology
- Search and navigation into a network of vocabularies
- Mapping of the terminology with a thesaurus
- Enrichment of a thesaurus
- Collaborative moderation of an update of the thesaurus



Terminology Management Platform (TMP)

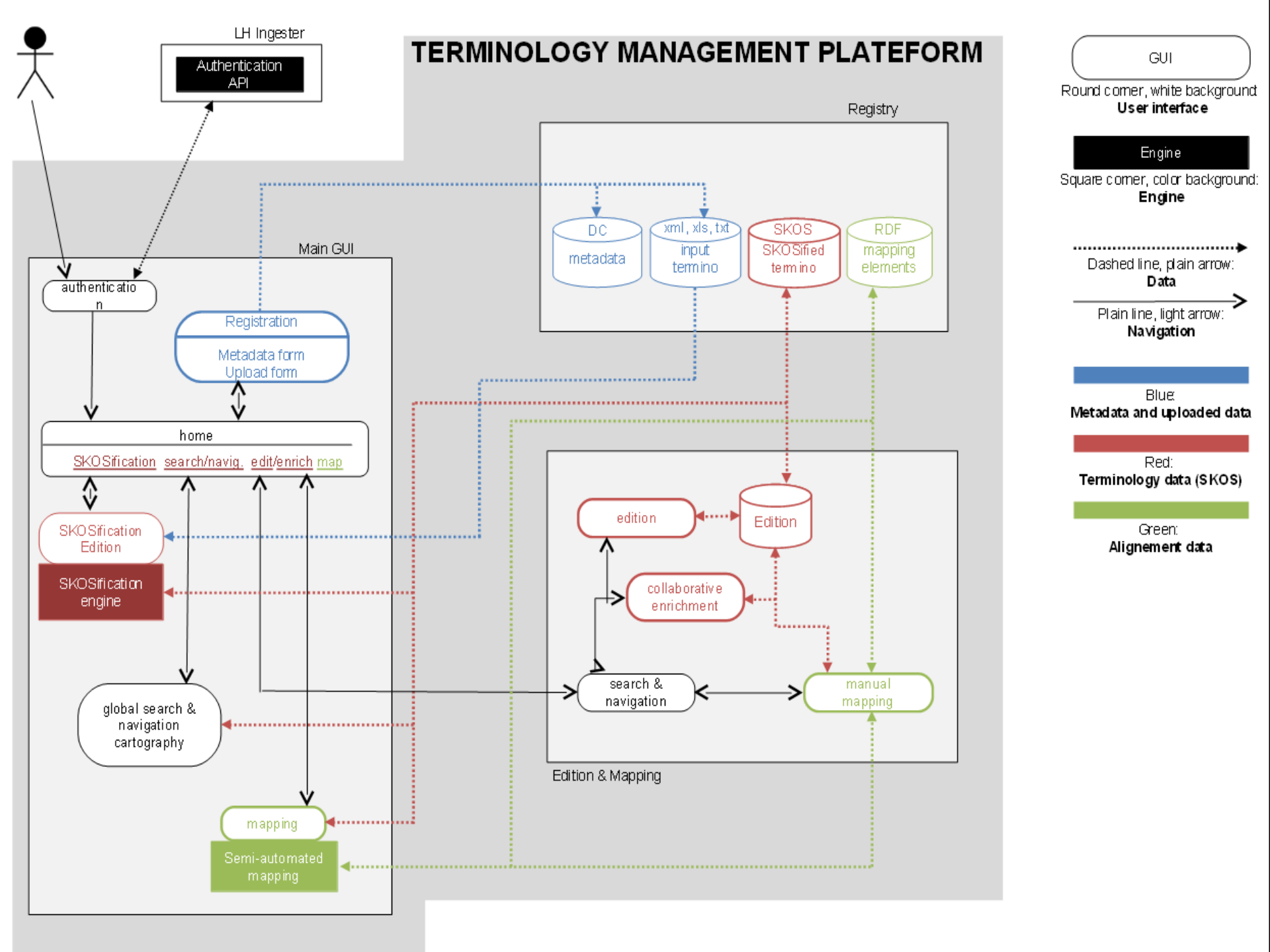
- To be a web service: For collaborative work online ;

- To have a user-friendly GUI: Adapted for a non-expert use in European museums, lirabries and archives

- To combine open-source components: Such a service must stay independent of proprietary codes and formats

- To be logically structured with an intuitive Workflow: The user must find which actions to do according to his/her needs

- To be flexible enough to be adapted to new standards: What if SKOS is updated in a new version or evolving towards an ontology description?


LH Ingester
Authentication API
TERMINOLOGY MANAGEMENT PLATEFORM
Registry
DC
metadata
xml, xls, txt
input
termino
SKOS
SKOSified
termino
RDF
mapping
elements
Main GUI
authentication
Registration
Metadata form
Upload form
home
SKOSification
search/navig.
edit/enrich
map
SKOSification
Edition
SKOSification
engine
global search &
navigation
cartography
mapping
Semi-automated
mapping
edition
Edition
collaborative
enrichment
search &
navigation
manual
mapping
Edition & Mapping
GUI
Round corner, white background
User interface
Engine
Square corner, color background:
Engine
Dashed line, plain arrow:
Data
Plain line, light arrow:
Navigation
Blue:
Metadata and uploaded data
Red:
Terminology data (SKOS)
Green:
Alignement data

Terminology Management Platform (TMP)

http://www.culture-terminology.org

TERMINOLOGY management PLATFORM

Home

TMP (Terminology Management Platform) is a "tool box" for creating, editing and managing thesaurus, classifications, subject headings, ontology and any other kind of terminology. This platform is mainly dedicated to cultural institutions from any sector (Libraries Archives and Museums) who hold or are willing to create terminologies.

You can register your terminology and proceed easily to the SKOSification (conversion into SKOS) and/or to the mapping of your terminology with another one.

Search public terminology

Search

Connect to TMP

Email

Password

Sign in! Forgot your password? Login

europeana think culture

LINKED HERITAGE

ATHENA Access to cultural heritage networks across Europe

Michael Multilingual Inventory of Cultural Heritage in Europe

Michael Culture Association

A project of LINKED HERITAGE Co-funded by the European Commission and the cip

Show 10 entries

Search:

Title	Subject	Languages	Author	
Being	Philosophy	fr	CR	View
Droit	Law	fr	CR	View
Fortification	Archaeology	de, en	JL	View
Furniture	Furniture	en, it	mvl	View
Objects	Economics	en, fr, en, fr	Mvl	View
ObjectsName	Archaeology	en, fr, nl	KMKG	View
Palissy	Furniture	fr	MCC-Inventaire	View
test	Decorative arts	ak	qdsfqsdf	View
titre	Decorative arts	fr	florent	View

Showing 1 to 9 of 9 entries

Previous Next

EN FR

TERMINOLOGY management PLATFORM

My home

Search therminologies

Search public terminologies my terminologies

Create new terminologies Create Import

Link two terminologies Mapping

Show 10 entries Search:

Title	Subject	Languages	Author	
Michael Subjects	Geography	en, fr, en, fr, en, fr, en, fr, en, fr	MCA	Edit export Delete
Objectnaam KIK	Archaeology	en, nl	VÃ©ronique	Edit export Delete
Ornamental Art	Decorative arts	hu	Julia Katona	Edit export Delete

TERMINOLOGY management PLATFORM

Create a terminology from scratch: fill in general information (metadata) on your terminology (* mandatory fields)

Title *		Aternative Title	
Subject *	Select a value...	Description	
Languages *	Select language...	Creation Date	
		Date of issue	
Edition *		Contributor	
Creator *		Identifier	
Publisher *		Source	
Rights *	Select a value	Amount of Concept	Select a value

Save Metadata

A project of LINKED HERITAGE Co-funded by the and the cip

1) Upload your terminology

2) Configure your import

Does your import look good ? No

Map you data on the underlying table with drop-down lists and when you mapping is OK click here.

This col is : Not mapped	This col is : Not mapped	This col is : Not mapped	This col is : Not mapped	This col is : Not mapped	This col is : Not mapped	This col is : Not mapped	This col is : Not mapped	This col is : Not mapped	T
ID	prefLabel@fra	prefLabel@nl	prefLabel@eng	broader	SN@nl	SN@en	SN@fr		
41875	Nom de l'objet	Objectnaam	Object name						1 T
41877	Architecture	Architectuur	Architecture	41875	Topterm voor architecturale elementen en onderdelen van architecturale elementen				2
41879	Construction	Bouwwerk	Single built work	41877	Algemene benaming voor alle soorten van (in aanbouw zijnde) gebouwen				3
45278	Composant de bâtiment	Deel van een gebouw	Building division	41879	Wordt gebruikt voor het benoemen van afzonderlijke kamers of ruimten van een gebouw				4

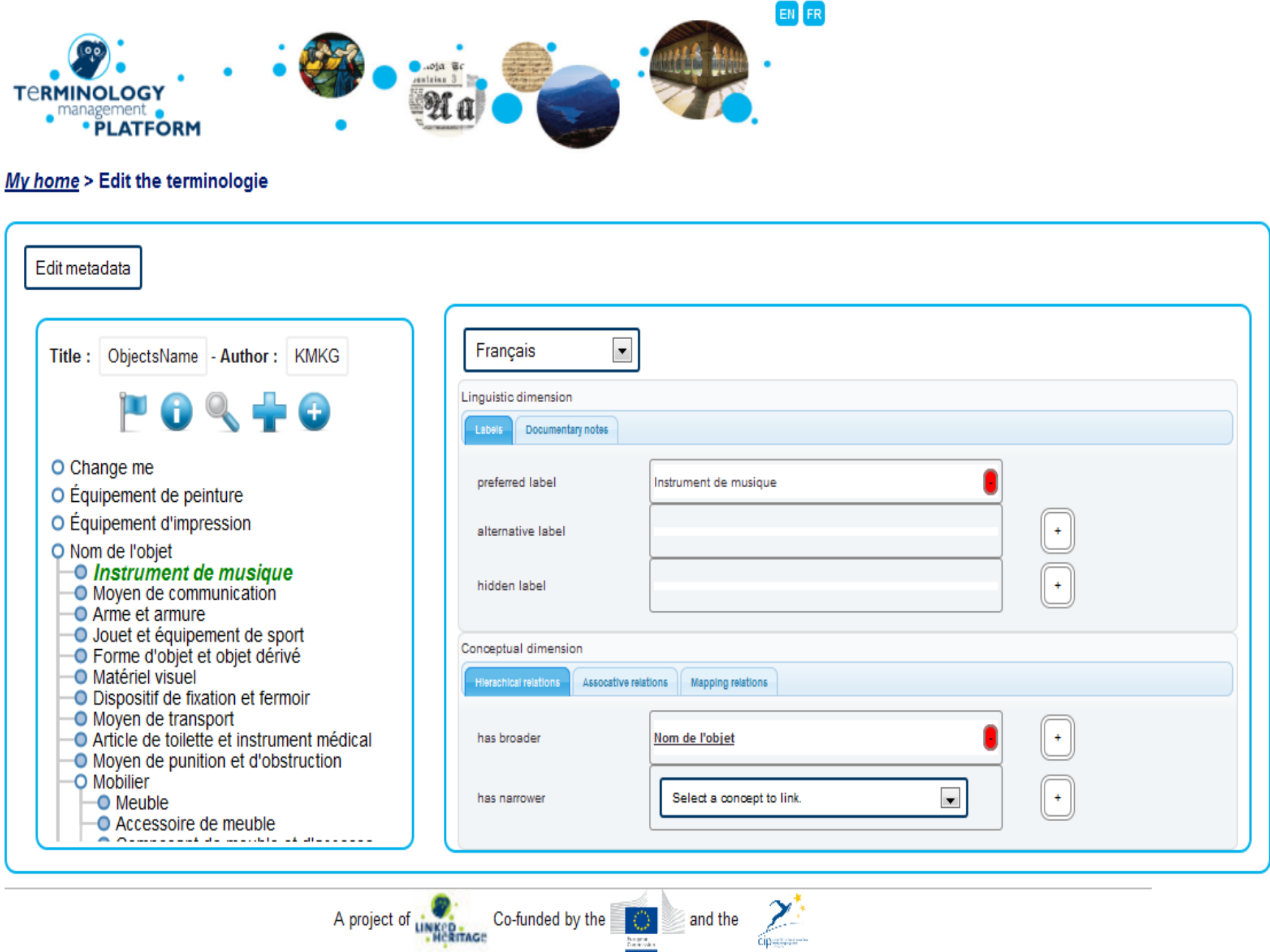
2) Configure your import

Does your import look good ? **No**

Map you data on the underlying table with drop-down lists and **when you mapping is OK click here.**

This col is : A local ID, Unique ?	This col is : A property, Of type: prefLabel, Language: French	This col is : A property, Of type: prefLabel, Language: Dutch	This col is : A property, Of type: prefLabel, Language: English	This col is : A reference, Of type: broader, To column: ID	This col is : A property, Of type: Select one	This col is : Not mapped
ID	prefLabel@fra	prefLabel@nl	prefLabel@eng	broader		SN@en
41875	Nom de l'objet	Objectnaam	Object name			
41877	Architecture	Architectuur	Architecture	41875		
41879	Construction	Bouwwerk	Single built work	41877		

Select one
Concept
altLabel
prefLabel
scopeNote
ConceptScheme
Collection
OrderedCollection
changeNote
definition
editorialNote
example
hiddenLabel
historyNote
notation
note


EN
FR
TERMINOLOGY
management
PLATFORM
My home > Edit the terminologie
Edit metadata
Title : ObjectsName - Author : KMKG
Change me
Équipement de peinture
Équipement d'impression
Nom de l'objet
Instrument de musique
Moyen de communication
Arme et armure
Jouet et équipement de sport
Forme d'objet et objet dérivé
Matériel visuel
Dispositif de fixation et fermoir
Moyen de transport
Article de toilette et instrument médical
Moyen de punition et d'obstruction
Mobilier
Meuble
Accessoire de meuble
Français
Linguistic dimension
Labels
Documentary notes
preferred label
Instrument de musique
alternative label
hidden label
Conceptual dimension
Hierachical relations
Assocative relations
Mapping relations
has broader
Nom de l'objet
has narrower
Select a concept to link.
A project of
Co-funded by the
and the

Title : PICO%20Thesaurus - **Author :** mibac

EUphoto

- when
- who
  - Persons
  - Corporate bodies
    - Families
    - Foreign Institutions
    - Performing art companies
    - Enterprises
    - Constitutional Bodies and Bodies of constitutional relevance
      - Senate of the Republic
      - National Council for Economics and Labour
      - Chamber of Deputies
      - Presidency of the Republic
      - Parliament
      - **Italian Government**
      - Court of Auditors
      - State Council
      - High Council of the Judiciary
      - Constitutional Court
    - Archives (institutions)
    - Independent Authorities
    - Museums
    - **Ministries, Departments, General Government Bodies**
    - Libraries, multimedia libraries and documentation centres
    - Defence and Security Bodies
    - Regional and Local Authorities
    - Religious Bodies
    - Publishing houses
    - Education and research institutions
    - Italian diplomatic and consular missions

Title : Euphoto - **Author :**

- types / photographic practice
- Photographic techniques
- keywords
  - health
  - society
  - religion and belief
  - science and technology
  - landscapes
  - politics
    - historical/political periods
    - fundamental rights
    - political events
    - government
      - *national government*
      - law enforcement
      - palace
      - parliament
      - political party
      - politician
      - regional government
      - public administration
      - local government
      - minister
      - heads of state
  - lifestyle and leisure
  - conflicts, war and peace
  - sport
  - human interest
  - education
  - labour

Select the type of the link

Select a value

closeMatch

Labels	Notes	Relations
preferred label	Italian Government	
alternative label		
hidden label		

Labels	Notes	Relations
preferred label	national government	
alternative label		
hidden label		

ATHENA PLUS
Access to cultural heritage networks for Europeana

Contact

Thank you for your attention !

Marie-Véronique Leroi (MCC)
marie-veronique.leroi@culture.gouv.fr

Eva Coudyzer (KMKG)
e.coudyzer@kmkg-mrah.be

www.athenaplus.eu

Twitter | https://twitter.com/AthenaPlusEU
LinkedIn | http://linkd.in/19XwWOz