

Michael Culture
Association



Recommendations and Guidelines for Terminologies



Marie-Véronique Leroi
Ministry of Culture and Communication
(MCC)
France

&

Eva Coudyzer
Royal Museum of Art and History
(KMKG)
Belgium

- **Presentation and Linked Heritage: work package overview**
- **Terminology ?**
- **The role of thesauri in information systems**
- **Athena legacy**
- **Your terminology as part of the semantic web**
- **Linked Heritage tools and work on terminologies**

- **Presentation and Linked Heritage work package overview**
- Terminology ?
- The role of thesauri in information systems
- Athena legacy
- Your terminology as part of the Semantic Web
- Linked Heritage tools and work on terminologies

- **Département de la Recherche, de l'Enseignement supérieur et de la technologie (DREST)**
 - **Service de la coordination des politiques culturelles et de l'innovation**
 - **Sécretariat général**

- **DREST is supported by Dédale (agency for culture, innovation and technology)**

- **DREST is mainly in charge of:**
 - **the national digitisation plan for cultural heritage**
 - **The representation of the ministry of culture in national authorities and at European level (MSEG, European projects such as Michael, Minerva, Europeana, Athena, Linked Heritage, ...)**
 - **The steering of the national inventory of digitised collections (“Patrimoine numérique”, national instance of Michael portal)**

- **As vice-president of the Michael Culture Association supporting the Minerva-network on behalf of the MCC, DREST takes an active part in European projects coordinated by the Mibac (2008-2011)**

■ **The KMKG-MRAH is a bilingual (French/Dutch) scientific research institute that operates under the Federal Belgian Science Policy. It has following missions (among others):**

- **Scientific research**
- **Conservation and restoration of museum collections**
- **Management and publication of collections through the development of:**
 - **a central database (MuseumPlus)**
 - **a collection publishing portal (www.carmentis.be)**
 - **metadata schemes and object descriptions by using standards**
 - **a data aggregator environment**
 - **a network of expertise in technological and content-based aspects of digitisation**

■ **Involvement in numerous projects stimulating access to cultural heritage, such as European CIP-projects (Athena, Linked Heritage etc.), Belgian science policy funded projects (MULTITA) etc.**

- **2011-2013 ; <http://www.linkedheritage.org>**
- **WP1: Project management and Coordination – Italy (ICCU)**
- **WP2: Linking Cultural Heritage Information – UK (CT)**
- ***WP3: Terminology - Belgium (KMKG)
& France (MCC)***
- **WP4: Public Private Partnership - UK (EDITEUR)**
- **WP5: Technical Integration - Greece (NTUA)**
- **WP6: Coordination of Content - Cyprus (CREF CYI)**
- **WP7: Dissemination & Training - Italy (UNIPD)**

■ WP-Leader: KMKG-MRAH, MCC

■ Technical partner:

- Université de Savoie (France),
- Instituto Superior Técnico Lissabon (Portugal),
- National Technical University of Athens (NTUA),
- DigiCULT (Germany)

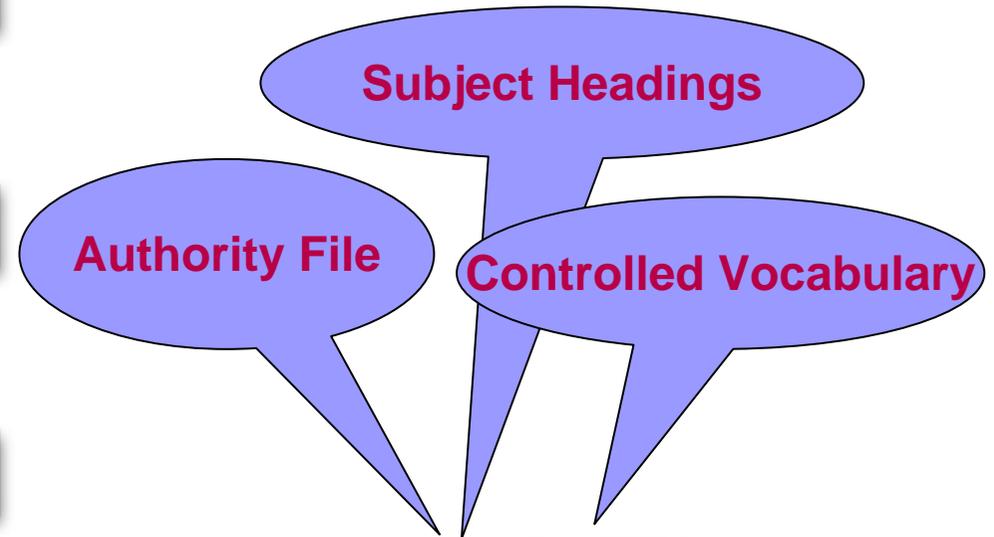
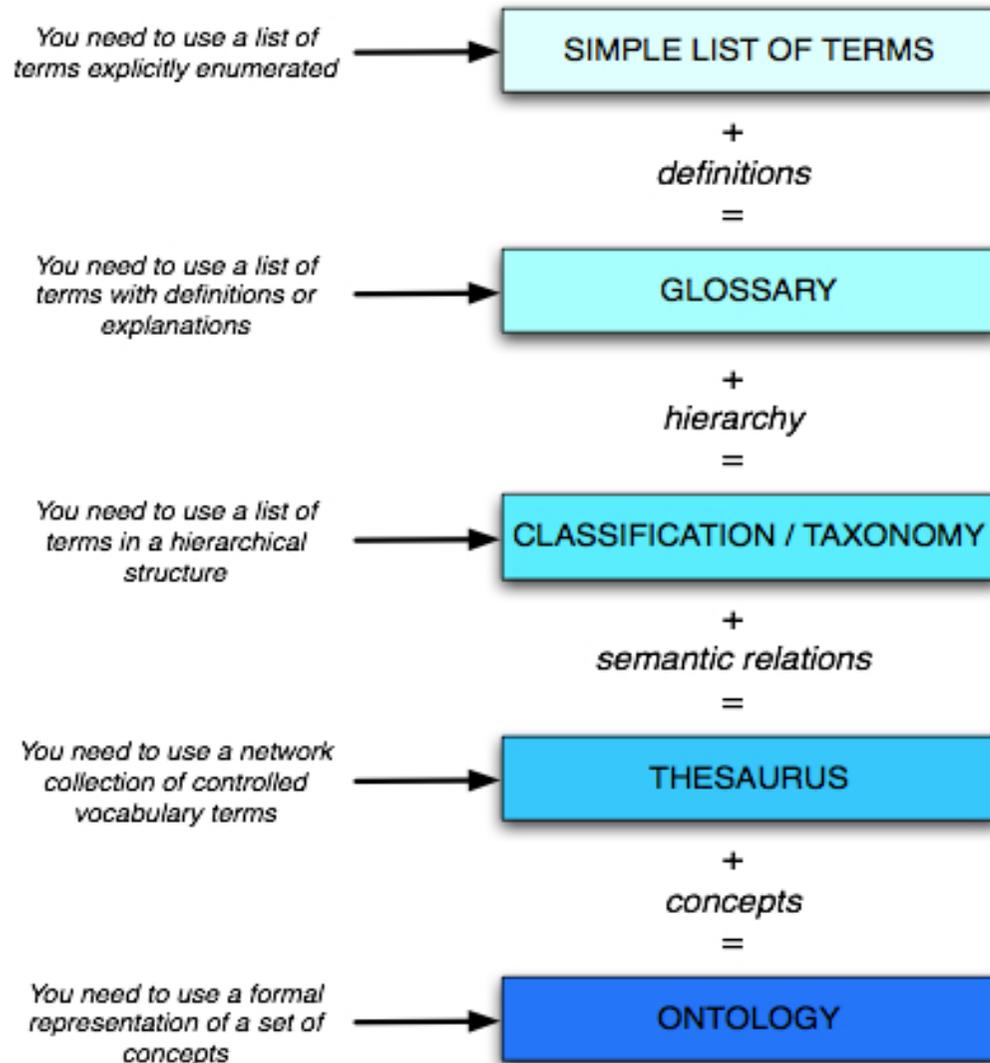
■ Aims to reduce the gap in terminology management skills that exists in cultural heritage institutions and the private sector by providing:

- Tools
- Guidelines
- Recommendations
- Training / workshops to be able to share complex and interoperable terminology resources with the community

■ The work of Linked Heritage will continue in EU-best practice network AthenaPlus (2013-2015)

- Presentation and Linked Heritage work package overview
- **Terminology ?**
- The role of thesauri in information systems
- Athena legacy
- Your terminology as part of the Semantic Web
- Linked Heritage tools and work on terminologies

Terminology ?



Terminology ?

- **Galleries, libraries, archives and museums (GLAM) use monolingual or multilingual terminologies.**
- **These terminologies can be: Thesauri, Classifications, Unstructured term lists, Taxonomies, Ontologies, etc.**
- **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”.**

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:

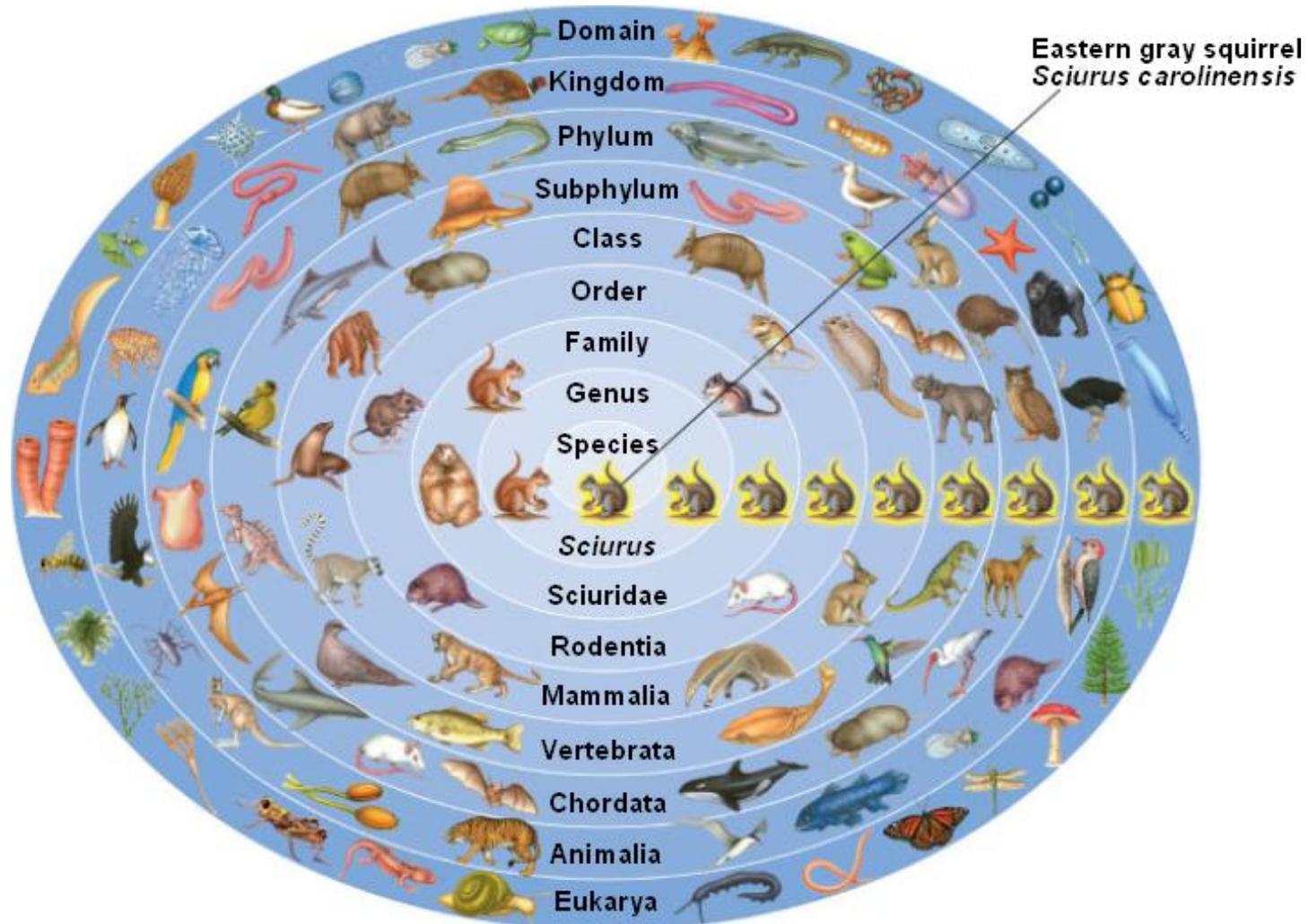
 World (facet)
 Europe (continent) (P)
 France (nation) (P)
 Île-de-France (region) (P)
 Ville de Paris, Département de (department) (P)
 Paris (inhabited place) (P)

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 : Linnean taxonomy of organisms



Some examples : Hornbostel-Sachs classification

Simple chordophones or zithers (31)

Instruments which are in essence simply a string or strings and a string bearer. These instruments may have a resonator box, but removing it should not render the instrument unplayable (although it may result in quite a different sound being produced). They include the piano therefore, as well as other kinds of zithers such as the koto, and musical bows.

Bar or stick zithers (311)

The string bearer is bar shaped.

- 311.1 Musical bows – The string bearer is flexible (and curved).
 - 311.11 Idiochord musical bows – The string is cut from the bark of the cane, remaining attached at each end.
 - 311.111 Mono-idiochord musical bows – Containing one string only
 - 311.112 Poly-idiochord musical bows or harp-bows – Containing several strings that pass over some type of bridge.
 - 311.12 Heterochord musical bows – The string is of separate material from the bearer.
 - 311.121 Mono-heterochord musical bows – The bow has one heterochord string only.
 - 311.121.1 Without resonator.
 - 311.121.11 Without tuning noose.
 - 311.121.12 With tuning noose.
 - 311.121.2 With resonator.
 - 311.121.21 With independent resonator.
 - 311.121.22 With resonator attached.
 - 311.121.221 Without tuning noose.
 - 311.121.222 With tuning noose.
 - 311.122 Poly-heterochord musical bows – The bow has several heterochord strings.
 - 311.122.1 Without tuning noose.
 - 311.122.2 With tuning noose.
 - 311.2 Stick zithers – With rigid string carrier
 - 311.21 Musical bow/stick – The string carrier has one rigid and one flexible end.
 - 311.22 True stick zithers – NB Round sticks which happen to be hollow by chance do not belong on this account to the tube zithers, but are round-bar zithers; however, instruments in which a tubular cavity is employed as a true resonator, like the modern Mexican harpa, are tube zithers.
 - 311.221 With one resonator gourd.
 - 311.222 With several resonator gourds.

- Presentation and Linked Heritage work package overview
- Terminology ?
- **The role of thesauri in information systems**
- Athena legacy
- Your terminology as part of the Semantic Web
- Linked Heritage tools and work on terminologies

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 role of thesauri in information environments



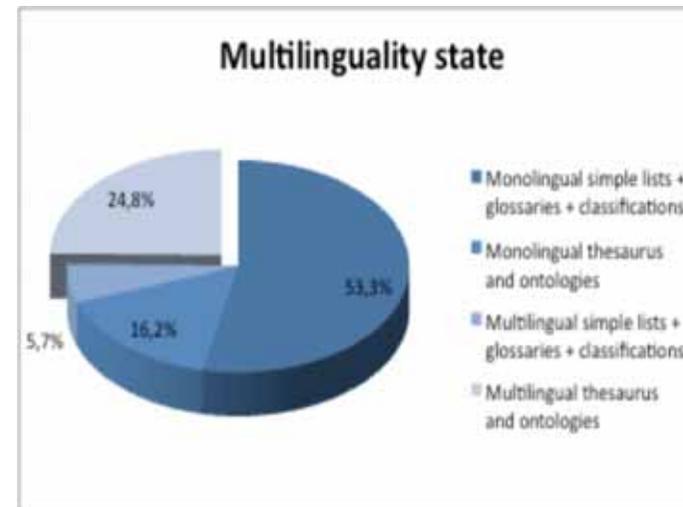
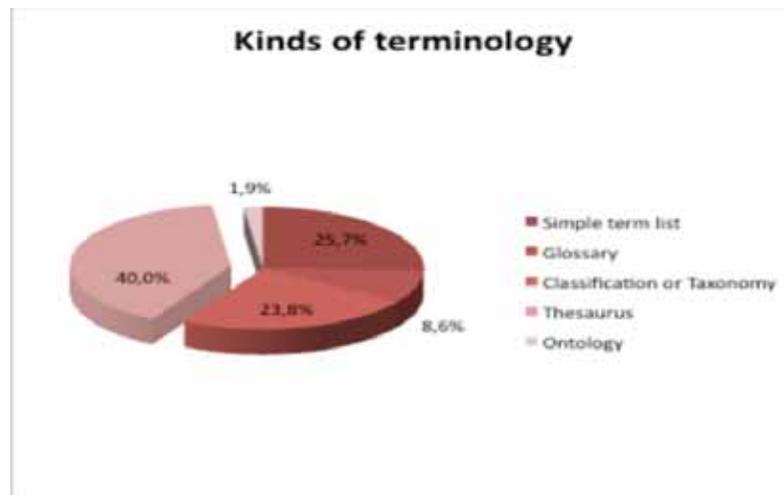
- **The colossal growth of information resources, demanding better subject identification**
- **The migration of traditional information resources to the Web, calling for more consistent subject approaches**
- **An urgent need for resource description and discovery through reuse of existing information management tool such as controlled vocabularies**
- **Problems assigned with the quality of unstructured information retrieved from the Web**
- **The need to provide users with knowledge structures such as thesauri for rapid and easy access to better-organized information (Ali Shiri)**

- Presentation and Linked Heritage work package overview
- Terminology ?
- The role of thesauri in information systems
- **Athena legacy**
- Your terminology as part of the Semantic Web
- Linked Heritage tools and work on terminologies

Athena benchmark study on terminology resources



- eContent Plus project : best practice network (2008-2011)
- EU Athena benchmark study (2009)
- Questions asked: Contact Information; Organisation's website (particularly any multilinguality) ; Detailed information about the terminology ; Use of the terminology ; Multilinguality of the terminology ; Availability of the terminology ; Audience for the terminology



Athena benchmark study on terminology resources



■ 44 terminologies were identified from 24 European countries

Terminology type	Number (of total 44)	%
Simple term list	3	7
Glossary	1	2
Classification / Taxonomy	11	24
Thesaurus	29	67
Ontology	0	0



■ Monolingual thesauri (66%)

- **Thesaurus iconographique : système descriptif des représentations (French)**
- **Terminology from national archives: topographic thesaurus (Swedish)**
- **Library Congress Subject Headings (LCSH) (English)**
- **Authority database of The National Library of Latvia (Latvian)**
- **Möbeltypologie (Westfälisches Museumsamt / Landesstelle für die nichtstaatlichen Museen in Bayern) (German)**



■ Multilingual thesauri (34%)

- **PACTOLS-thesaurus (peuples et cultures, anthroponymes, chronologie relative, toponymes, oeuvres, lieux, sujets) (French, English, Italian, German, Spanish – near future: Dutch).**
- **The Multilingual Egyptological Thesaurus (English, Dutch, German, French, Italian, Spanish And Portuguese)**
- **UNESCO thesaurus (English, French, Spanish, Russian)**
- **GEMET, the General Multilingual Environmental Thesaurus- European Environment Agency (Basque, Bulgarian, Dutch, Finnish, French, German, Hungarian, Italian, Norwegian, Portuguese, Russian, Slovenian, Spanish, Danish, Slovak, Swedish, Greek)**

- Presentation and Linked Heritage work package overview
- Terminology ?
- The role of thesauri in information systems
- Athena legacy
- **Your terminology as part of the Semantic Web**
- Linked Heritage tools and work on terminologies

- **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

“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)

How to join the Semantic Web?



■ Enhance semantic interoperability

- **Use/create terminology type: thesaurus**
- **Preferably multilingual**
- **Imply international thesauri-standards (e.g. ISO-norms)**
- **Conceptualise thesaurus terms (by means of Uniform Resource Identifiers (URI) data can be connected that was not associated before**

■ Enhance technical interoperability

■ Use standardised formats:

- **Extensible Markup Language (XML)**
- **Resource Description Framework (RDF)**
- **Web Ontology Language (OWL)**
- **Simple Knowledge Organisation System (SKOS)**

Athena/Linked Heritage Booklet



- **Recommendations for design and management (2011):**
<http://www.linkedheritage.eu/getFile.php?id=244>

"TERMINOLOGY" COVERS SEVERAL TYPES OF VOCABULARIES:
LEXICON, THESAURUS, CONTROLLED VOCABULARY, ONTOLOGY...

THE TYPE OF VOCABULARY IS HIGHLY CONNECTED TO
ITS PURPOSE (INDEXATION, SEARCH AND RETRIEVAL,
TRANSLATION, ...). SKOSIFYING YOUR TERMINOLOGY IS
A CRUCIAL STEP TO SHARE IT AND TO CONNECT IT TO
OTHER TERMINOLOGY RESOURCES AND TO IMPROVE
MULTILINGUALISM AT EUROPEAN LEVEL.

www.athenaeurope.org/athenawiki



<http://www.linkedheritage.org> | info@linkedheritage.org



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

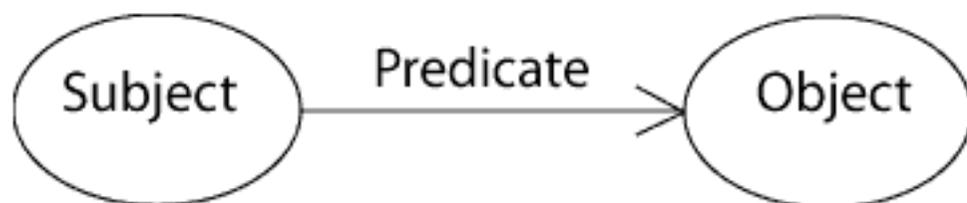


SKOS : Something Kool Original and Sexy



- **Simplified Knowledge Organisation System (SKOS)**
- **Recommendation of the W3C since August 2009**
- **Interoperable format for thesaurus**
- **Descriptors and non-descriptors became SKOS concepts expressed with labels**
- **Ideal compromise between a thesaurus and an ontology (RDF/XML)**

■ RDF : Resource Description Framework

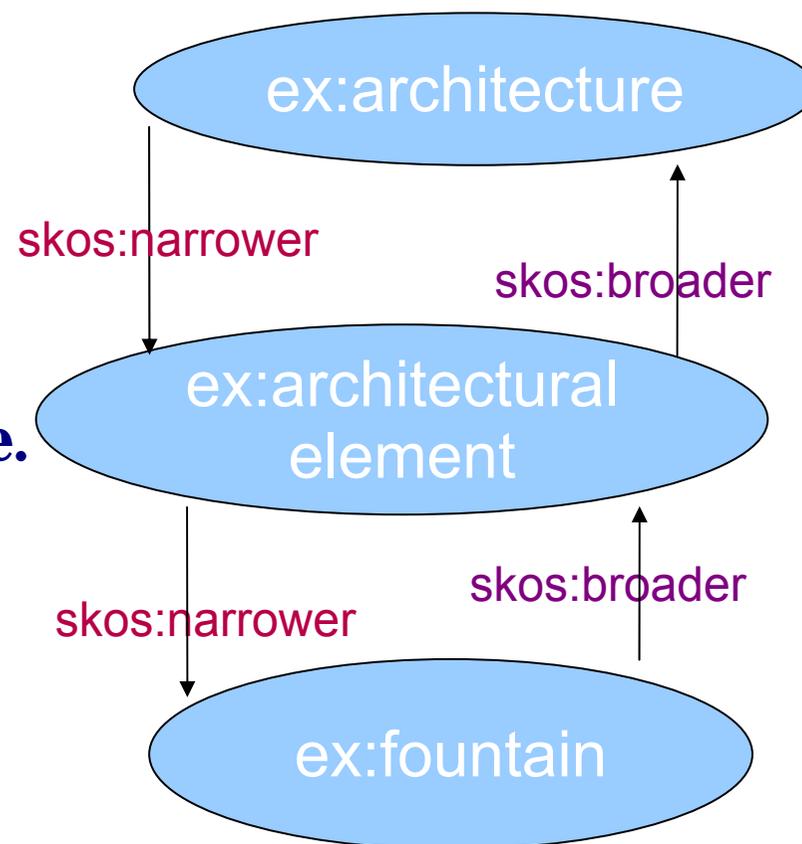


■ Each information is structured as a triple.

■ Subjects and objects have an URI

■ URI = Unique identifier

■ URI = one of the Semantic Web & Linked Data principle



- **Concept : unit of thought (*skos:Concept*)**
- **Concept scheme : groups of concepts, micro-thesaurus (*skos:ConceptScheme*)**
- **Collection : thematic groups of concepts (*skos:Collection*)**
- **Concepts are expressed with labels :**
 - **Preferred Label : *skos:prefLabel***
 - **Alternative label : *skos:altLabel***
 - **Hidden Label : *skos:hiddenLabel***
- **Semantic relations :**
 - **Hierarchical : *skos:broader* / *skos:narrower***
 - **Associative : *skos:related***



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

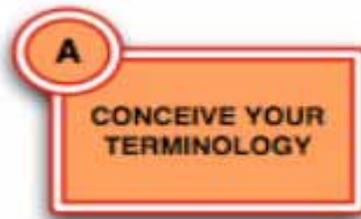
Create or adapt an existing terminology

■ **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)

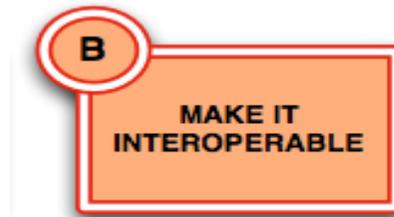
■ **C : How can I link my terminology to other terminologies?**

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



□ Conceive a terminology : Create or adapt an existing terminology

- **Define your collection domain(s)**
- **Identify your users' expectations (about your semantic descriptions)**
- **Define your connection with the datamodel**
- **Choose the terms for the semantic description of your digital resources**
- **Organise your terms into a thesaurus structure**
- **Find equivalent terms in other languages**
- **Implement your thesaurus**



- ❑ **Make the terminology interoperable : Technical implementation (=> Semantic Web)**
 - **SKOS: Simple Knowledge Organization System**
 - **Standard created by the World Wide Web Consortium**
 - **Allows flexible and machine-readable encoding of thesauri / other controlled vocabularies**
 - **Within this standard thesaurus relationships can be encoded in such a way as to be used by various web services and web research systems**
 - **A large number of terminologies will be brought in SKOS, to flexibly and consistently function within the semantic web environment**



- **Link the terminology to a network : Linking with other resources (=> Linked Data)**
 - **Links several different thesauri**
 - **Expand search functionalities through federated searching of multiple controlled vocabularies and linked data sources**
 - **Allow for the integration of thesauri into many web-based search engines and services**
 - **Provide semantically rich visualisation of thesauri and links between and among thesauri**
 - **Facilitate multilingual information access and retrieval**
 - **Provide easy access to thesauri for indexing and information representation purposes (Ali Shiri)**

- **« Controlled and structured vocabulary in which concepts are represented by terms, organized so that relationships between concepts are made explicit, and preferred terms are accompanied by lead-in entries for synonyms or quasi-synonyms » (ISO 25964-1)**
- **Norm ISO 25964-1 published in August 2011 : Information and documentation -- Thesauri and interoperability with other vocabularies -- Part 1: Thesauri for information retrieval**
- **Second part on Interoperability to be published in 2012**
- **Revision of the ISO 2788, ISO 5964 and BS8723**
- **Thesaurus oriented and evolution of SKOS**

- Presentation and Linked Heritage work package overview
- Terminology ?
- The role of thesauri in information systems
- Athena legacy
- Your terminology as part of the Semantic Web
- **Linked Heritage tools and work on terminologies**

■ Mapping experiments with:

- **UK: “Vocabulary BT”: British Museum thesaurus, 7245 concepts, in English**
- **BE: “Vocabulary KMKG Nom de l’objet thesaurus”: 2469 concepts, in French, Dutch and English, mapped to British Museum Thesaurus**
- **IT: “Vocabulary ICCD nom de l’objet thesaurus”: 99 concepts, in Italian**
- **IL: “Vocabulary Israel Museum”: imported in test, 4070 concepts, in Hebrew and English**
- **BE: “Vocabulary Museum Plantin-Moretus Prentencabinet”: 119 concepts (more concepts still to be added), in Dutch**
- **BE: “Vocabulary Stedelijke Musea Mechelen” created**
- **DE: “Vocabulary möbeltypologie”, in German**

- **LIDO (Lightweight Information Describing Objects)**
- **Provides an explicit format to deliver (museum's) object information in a standardized way**
- **XML Schema for Contributing Content to Cultural Heritage Repositories:**
 - **for delivering metadata, for use in a variety of online services, from an organization's online collections**
 - **database to portals of aggregated resources – as well as exposing, sharing and connecting data on the Web.**
 - **Intended to represent the full range of descriptive information about museum objects, e.g. art, cultural, technology and natural science.**
 - **It supports multilingual environments.**

LIDO actor role: "art historian"



Language	MALE SINGULAR	FEMALE SINGULAR	PLURAL	ALTERNATIVE LABELS
Croatian	Povjesničar umjetnosti	Povjesničarka umjetnosti	Povjesničari umjetnosti	
Czech	Kunsthistorik	Kunsthistorička	Kunsthistorici, Kunsthistoričky	Historik umění, historička umění
Dutch	Kunsthistoricus	Kunsthistorica	Kunsthistorici, Kunsthistorica's	
English	Art historian		Art historians	
French	Historien d'art	Historienne d'art	Historiens d'art, historiennes d'art	
Greek	Istorikos tis technis	Istorikos tis technis	Istoriki tis technis	
Hungarian	Művészettörténész	Művészettörténész	Művészettörténészek	
Hebrew	חוקר תולדות האמנות	חוקרת תולדות האמנות	חוקרי תולדות האמנות	
Italian	Storico dell'arte	Storica dell'arte	Storici dell'arte, Storiche dell'arte	
Polish	Historyk sztuki	Historyczka sztuki	Historycy sztuki	
Slovak	Kunsthistorik	Kunsthistorička	Kunsthistorici, Kunsthistoričky	Historik umenia, historička umenia
Portugese	Historiador de arte	Historiadora de arte	Historiadores de arte, Historiadoras de arte	Historiadoras de arte
Romanian	Istoric de artă		Istorici de artă	
Slovenian	Umetnostni zgodovinar	Umetnostna zgodovinarica	Umetnostni zgodovinarji, Umetnostne zgodovinarke	strokovnjak za umetnostno zgodovino, strokovnjakinja za umetnostno zgodovino
Spanish	historiador de arte	historiadora de arte	historiadores de arte, historiadoras de arte	historiadoras de arte
Swedish	Konsthistoriker	Konsthistoriker	Konsthistoriker, Konsthistoriker	

- **Shiri A., Powering search. The Role of Thesauri in New Information Environments, New Jersey 2012**
- **Aitchison J., Gilchrist A., Bawden D., Thesaurus Construction and use: A Practical Manual, 4th. ed. London 2000**
- **Hodge G., Systems of knowledge organization for digital libraries: Beyond traditional authority files, Washington 2012.**
- **ISO 25964-2, Information and documentation - Thesauri and interoperability with other vocabularies -- Part 2: Interoperability with other vocabularies, Geneva, March 2013**
- **BS 8723: Structured Vocabularies for Information Retrieval, London 2005**
- **ANSI/NISO Z39: Guidelines for the construction, format, and management of monolingual controlled vocabularies, Bethesda, 2005**
- **SKOS: <http://www.w3.org/2004/02/skos/>**

Merci!

marie-veronique.leroi@culture.gouv.fr

e.coudyzer@kmkg-mrah.be

<http://www.linkedheritage.org/>



ATHENA
Access to cultural heritage
networks across Europe

