

SSHOC WP updates on vocabularies

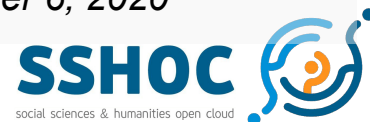


1. Vocabulary Survey by **Clara Petitfils & Nicolas Larrousse (WP7, WP3)**
2. Vocabulary mapping tool for archaeology in ARIADNEplus by **Holly Wright (WP5)**
3. Management of controlled vocabularies within the Aioli platform with the Opentheso tool by **Isabelle Cao (WP4)**
4. BackBone Thesaurus, a model for thesauri interoperability by **Chryssoula Bekiari (WP4)**
5. How Dataverse supports external vocabularies by **Slava Tykhonov (WP5)**

SSHOC Considerations for Vocabulary Platforms, virtual workshop, November 6, 2020



This project is funded from the EU Horizon 2020 Research and Innovation Programme (2014-2020) under Grant Agreement No. 823782



T7.4's survey on the use of SSH vocabularies in the research community

T7.4 – Governance : Population, Curation & Sustainability of the SSH Open Marketplace

D7.6 – Resources for Marketplace content description

Nicolas Larrousse & Clara Petitfils
CNRS/Huma-Num

**Online workshop – “SSHOC Considerations for the Vocabulary Platforms”
6th November 2020**

The Survey

The CLARIN logo is located in the top right corner. It features a stylized graphic of a network or molecular structure above the word "CLARIN" in a bold, sans-serif font.

- D7.6 Resources for Marketplace content description due for M24, 31/12/2020
- From February to May 2020 on LimeSurvey
- 13 questions organised into 3 sets :

(1) Respondent's profile : SSH discipline, research organisation, country of residence

(2) Use of SSH vocabularies : which vocabularies, alignment, language, online availability, outdatedness

(3) Participation on a follow-up group relating to this survey

- Objectives of this survey :

(1) Conduct an analysis on the use of vocabularies within the SSH research communities

(2) Support the development of the SSH Open Marketplace through its mapping process

(3) Provide a complementary analysis regarding other partners & projects

The results



CLARIN

- **330 answers but that need to be differentiated between :**

→ Incomplete answers : when respondents didn't complete the survey entirely → **258**

> 39% stopped at the first set of questions and 89% at the very beginning of the second set, the fourth question

→ Complete answers : completed from the first question to the last one → **72**

- **Question 4 « Do you / your organisation use any vocabulary(ies) ? Both digital and printed ones »**

→ Out of the 72 complete answers à **52** "YES" answers with 90% of them indicating which one they use

The results

The CLARIN logo is located in the top right corner. It features a stylized graphic of a network or molecular structure with blue dots and lines, positioned above the word "CLARIN" in a bold, blue, sans-serif font.

- **Respondents' profile :**

- 3 questions

- In D7.6, distinction between all answers considered, incomplete and complete ones

- **SSH discipline :**

- Out of all answers (330) :

- > 15% Linguistics

- > 11,4% Archaeology & Prehistory

- > 10,8% for Sociology

- > 7,8% for History

The results

- **Respondents' organisation type :**

- 32,9% from Universities & Research Performing Organisations
- 20% from Universities
- 3,9% from Research Libraries & Archives
- 1,5% from Research & e-Infrastructures and EOSC thematic clusters

- **Respondents' country of residence :**

- 38,8% from France
- 2,7% from Germany
- 2,7% from The Netherlands



CLARIN

The results

The CLARIN logo is located in the top right corner. It features a stylized graphic of a network or molecular structure with blue dots and lines, positioned above the word "CLARIN" in a bold, blue, sans-serif font.

- **Q4 : 159 entries completed**

→ None “massively” used by the community, but some repeatedly mentioned : DDI, Getty, CESSDA Controlled Vocabularies, ELSST, Dublin Core, Pactols

→ Important to note that the understanding of what a vocabulary is, differs a lot from one researcher to another

- **// Q9 “With which vocabulary(ies) would you like the one(s) you’re using to be aligned with?”**

→ Getty & Wikidata.

- **Q6 “In which languages is/are this/those vocabularies available in?”**

→ (1) English, (2) French and (3) German & Spanish

French bias : because of the large amount of French residents respondents

Links with the SSH Open Marketplace

The CLARIN logo features a stylized network of blue dots connected by lines, with the word "CLARIN" in blue capital letters below it.

- **Vocabularies are crucial for the working of the SSH Open Marketplace as a discovery platform.**
- **Two potential use of vocabularies**
 - Integration if it's seems relevant into the Marketplace to describe entries and thus foster discoverability
 - Consider it as a resource and create an entry in the Marketplace

Links with the SSH Open Marketplace

The CLARIN logo features a stylized network of blue dots connected by lines, with the word "CLARIN" in blue capital letters below it.

- Challenges of integration as a resource to describe items in the Marketplace

→ The main part of the content of the Marketplace comes from “ingestion” of external sources with “local” vocabularies

→ Need to set up a mapping process which could be complex: intertwining machine and human action

Thank you for your attention!

To contact us :

clara.petitfils@huma-num.fr

nicolas.larrousse@huma-num.fr

Vocabulary mapping tool for archaeology in **ARIADNE***plus*



Holly Wright, Archaeology Data Service

Ceri Binding & Douglas Tudhope

University of South Wales, Trefforest

[tgn:7029392](#) World

[tgn:1000003](#) Europe

[tgn:7008591](#) United Kingdom

[tgn:7002443](#) Wales

[tgn:7018963](#) Rhondda Cynon Taf

[tgn:7441565](#) Trefforest



- **ARIADNE**

- 24 partners, 13 countries, 9 languages, 27 subject vocabularies
- 1.9 million data records aggregated/integrated
- Subject vocabularies coordinated via mapping to Getty AAT – total 6416 mappings produced

- **ARIADNE^{plus}**

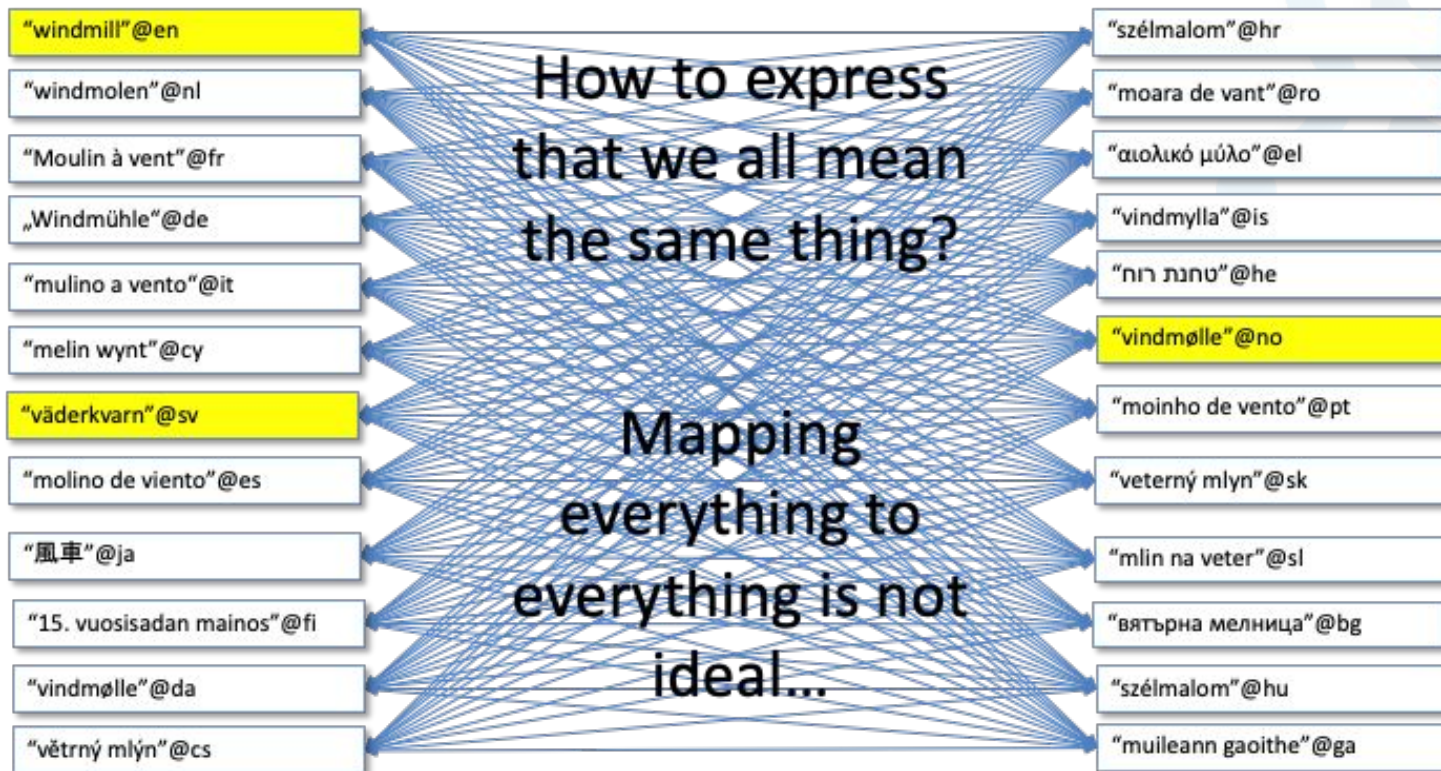
- 41 partners, 29 countries, 22 languages, ?? subject vocabularies
- Data aggregation/integration work currently in progress
- Reusing, revising and supplementing previous mappings
- Adding vocabulary mappings from new data partners
- Adding Wikidata mappings (multilingual entry vocabulary)

Why do we need vocabulary matching in ARIADNE?



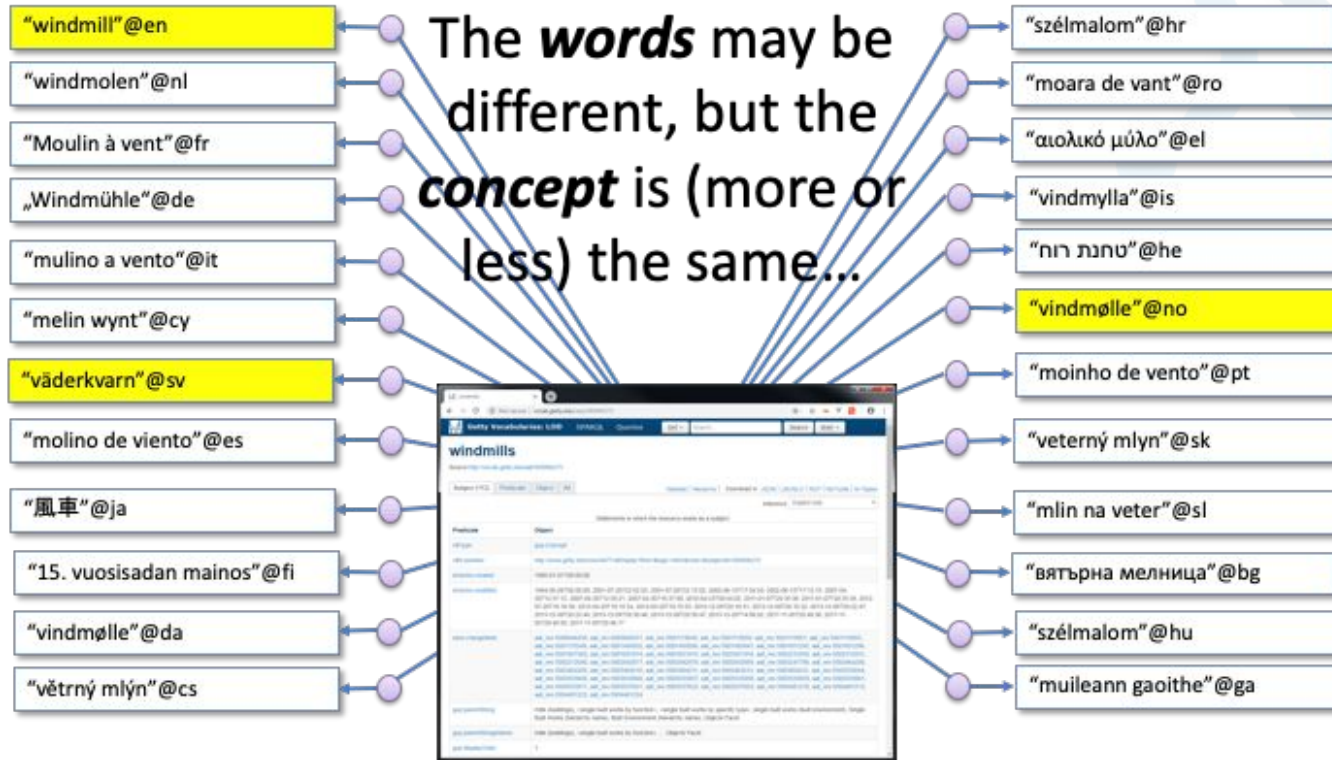
- Source datasets not necessarily produced with aggregation, consolidation, cross-search and reuse in mind
- I say "*potato*", you say "*pomme de terre*", she says "*maris piper*" – multiple barriers to cross-searching subject metadata: language, punctuation, spelling, homonyms, synonyms, level of specificity
- Text-based search is limited by any/all of these
- Need to establish mutually agreed meaning...

Multilingual subject index terms



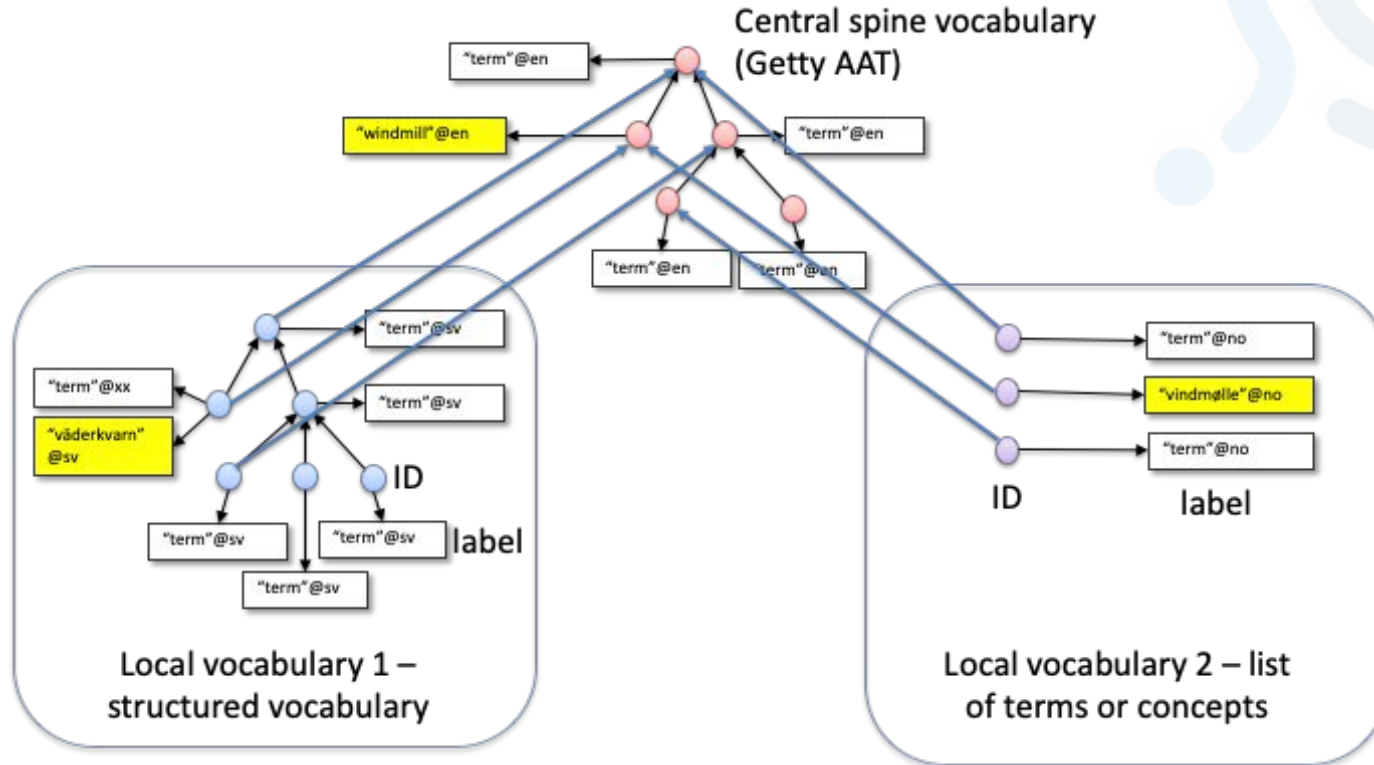
Ideally we want to include any/all of these variants in a single query

Map local terms to a central concept



Now we can include any/all of these variants in a single query

Map local concepts to a central spine



Vocabulary Matching Tool

<https://vmt.ariadne.d4science.org/vmt/>



- For matching local subject terms / concepts to Getty AAT concepts
- Search & browse Getty AAT
- No auto match: examine scope and context of source / target concepts

Vocabulary Matching Tool

English

Source Concept	Match Type	Target Concept	Suggest	Delete Row
Identifier		Filter column...		
Label		Filter column...		
https://purl.org/heritagedata/schemes... Abbey Church	Close Match	abbey churches	Q	
https://purl.org/heritagedata/schemes... ABBEY	Exact Match	abbeys (non-monasteries)	Q	
https://purl.org/heritagedata/schemes... AGRICULTURAL BUILDING	Exact Match	agricultural buildings	Q	
https://purl.org/heritagedata/schemes... AGRICULTURAL DWELLING	Broad Match	agricultural buildings	Q	
https://purl.org/heritagedata/schemes... AGRICULTURAL HALL	Broad Match	agricultural buildings	Q	
https://purl.org/heritagedata/schemes... FARM BUILDING	Close Match	agricultural buildings	Q	
https://purl.org/heritagedata/schemes... FIELD SYSTEM	Broad Match	agricultural land	Q	
https://purl.org/heritagedata/schemes... FIELD SYSTEM	Broad Match	agricultural land	Q	
https://purl.org/heritagedata/schemes... LAND USE SITE	Broad Match	agricultural land	Q	
https://purl.org/heritagedata/schemes... LYNCHET	Broad Match	agricultural land	Q	
https://purl.org/heritagedata/schemes... CURVILINEAR ENCLOSURE	Broad Match	agricultural settlements	Q	
https://purl.org/heritagedata/schemes... DITCHED ENCLOSURE	Broad Match	agricultural settlements	Q	
https://purl.org/heritagedata/schemes... DOUBLE DITCHED ENCLOSURE	Broad Match	agricultural settlements	Q	
https://purl.org/heritagedata/schemes... ENCLOSED SETTLEMENT	Broad Match	agricultural settlements	Q	
https://purl.org/heritagedata/schemes... ENCLOSURE	Broad Match	agricultural settlements	Q	
https://purl.org/heritagedata/schemes... AGRICULTURE AND SUBSISTENCE	Broad Match	agriculture	Q	
https://purl.org/heritagedata/schemes... AIR RAID SHELTER	Exact Match	air raid shelters	Q	
https://purl.org/heritagedata/schemes... AIRCRAFT	Close Match	aircraft	Q	

390 rows

IMPORT JSON EXPORT JSON EXPORT CSV ADD NEW ROW CLEAR ROWS SHOW HELP

ARIADNE plus

Created by University of South Wales

ARIADNEplus is a Horizon 2020 project funded by the European Commission (Grant Agreement No. 823782)

This application retrieves some information originating from Getty Art & Architecture Thesaurus (AAT)® which is made available under the ODC Attribution License. See <https://ocds.getty.edu/> for further details.

Type of match between concepts

Exact Match



BUT: don't rely on label matches;
consider full context – meaning and
scope of concepts

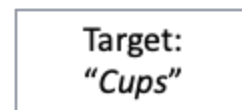
Close Match



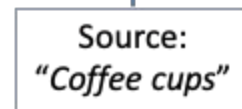
Where scope or context of concepts
suggests slight conceptual differences

[Note: `skos:narrowMatch` also exists]

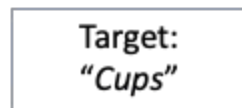
"Some/all" rule for generic
hierarchical relationships:



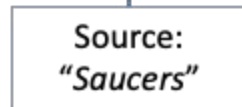
Broad Match



Some cups are
coffee cups; all
coffee cups are
cups



Related Match



Some other
association
exists between
the concepts.
Where possible
prefer one of
the other match
types though

Why? Using vocabulary mappings

- Search term = "CEMETERY"
 - may retrieve some results. Expand with plural "*cemeteries*". May retrieve a few more...
- What do the mappings give us?
 - local:12345 "CEMETERY" skos:exactMatch [aat:300266755](#) "*cemeteries*"
- What other local vocabulary terms are mapped to [aat:300266755](#) ?
 - "BARROW CEMETERY", "CHOLERA BURIAL GROUND", "FRIENDS BURIAL GROUND", "INHUMATION CEMETERY", "JEWISH CEMETERY", "MUSLIM CEMETERY", "NONCONFORMIST CEMETERY", "PLAGUE CEMETERY", "ROMAN CATHOLIC CEMETERY", "WALLED CEMETERY"
- We now have an expanded search, and have uncovered potential links between records indexed using any of these terms.
 - However they are all in one language...

Why? Using vocabulary mappings

- Multilingual terms associated with concept [aat:300266755](#) ?
 - **Preferred labels:** "cemeteries"@en, "campos santos"@es, "campi santi"@it, "cimetières"@fr, "begraafplaatsen"@nl, "Friedhof"@de
 - **Alternate labels:** "cemetery"@en, "campos santos (cemeteries)"@en, "campo santo (cemetery)"@en, "campo santo"@es, "campo santo"@it, "cimetière"@fr, "cœmeterium (cemeteries)"@la, "camposanto (cemetery)"@en, "camposanto"@it, "begraafplaats"@nl, "Friedhöfe"@de
- Can we utilize AAT (poly)hierarchical structure? (Yes!)
 - AAT concepts narrower (more specific) than *cemeteries*:
 - *catacombs, columbaria (cemeteries), graveyards, lawn cemeteries, memorial parks, necropolises, Reihengräberfelder, churchyards, cineraria (cemeteries), military cemeteries (veteran cemeteries), national cemeteries, pet cemeteries, potter's fields, war cemeteries*
 - Plus each of these concepts has multilingual preferred / alternate terms - we now have a semantically expanded multilingual search

Why? Using vocabulary mappings

- And finally...
 - Wikidata contains mappings to AAT concepts
 - [wikidata:Q39614](https://www.wikidata.org/wiki/Q39614) is already directly mapped to [aat:300266755](https://www.getty.edu/research/conducting_research/vocabularies/aat/aat300266755.html) (“cemeteries”) and has ***many*** more multilingual labels:
 - Cemetery, graveyard, burial ground, cemeteries, churchyard, cimetière, champ de repos, boulevard des allongés, champ du repos, Friedhof, Totenacker, Begräbnisplatz, Gottesacker, Kirchhof, Leichenhof, Begraafplaats, Asie, Fosal, Fosar, Zimenterio, Corralón, Fusal, Sagrero, Fosal d'os moros, Cimenterio, Fonsal, 墳場, cmentarz, cemitério, pokopališče, гробље etc.
- One mapping brings in many alternative terms/concepts to improve multilingual query experience and to expand potential results.
- Use of semantic links can improve recall without necessarily sacrificing precision

RDF serialisations of mappings

```
<?xml version="1.0" encoding="UTF-8"?>
<!--Example mappings expressed in RDF/XML serialization format (-->
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:skos="http://www.w3.org/2004/02/skos/core#"
  >
  <skos:Concept rdf:about="http://snd.gu.se/sv/catalogue/keyword/bengomma">
    <skos:prefLabel xml:lang="sv">bengömma</skos:prefLabel>
    <skos:closeMatch rdf:resource="http://vocab.getty.edu/aat/300265420"/><!--remains-->
  </skos:Concept>

  <skos:Concept rdf:about="http://snd.gu.se/sv/catalogue/keyword/bergshistorisk-lamning-ovrig">
    <skos:prefLabel xml:lang="sv">bergshistorisk lämning övrig</skos:prefLabel>
    <skos:closeMatch rdf:resource="http://vocab.getty.edu/aat/300006423"/><!--mine structures-->
  </skos:Concept>

  <skos:Concept rdf:about="http://snd.gu.se/sv/catalogue/keyword/bildristning">
    <skos:prefLabel xml:lang="sv">bildristning</skos:prefLabel>
    <skos:broadMatch rdf:resource="http://vocab.getty.edu/aat/300080131"/>
  </skos:Concept>

  <skos:Concept rdf:about="http://snd.gu.se/sv/catalogue/keyword/björngrav">
    <skos:prefLabel xml:lang="sv">björngrav</skos:prefLabel>
    <skos:broadMatch rdf:resource="http://vocab.getty.edu/aat/300005907"/>
  </skos:Concept>
</rdf:RDF>

# Mappings expressed in Turtle RDF serialization format
@prefix data: <http://snd.gu.se/sv/catalogue/keyword/> .
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .
@prefix aat: <http://vocab.getty.edu/aat/> .

data:bengomma a skos:Concept;
  skos:prefLabel "bengömma"@sv ;
  skos:closeMatch aat:300265420 . # remains

data:bergshistorisk-lamning-ovrig a skos:Concept ;
  skos:prefLabel "bergshistorisk lämning övrig"@sv ;
  skos:closeMatch aat:300006423 . # mine structures

data:bildristning a skos:Concept;
  skos:prefLabel "bildristning"@sv ;
  skos:broadMatch aat:300080131 . # rock carvings

data:björngrav a skos:Concept;
  skos:prefLabel "björngrav"@sv ;
  skos:broadMatch aat:300005907 . # graves
```

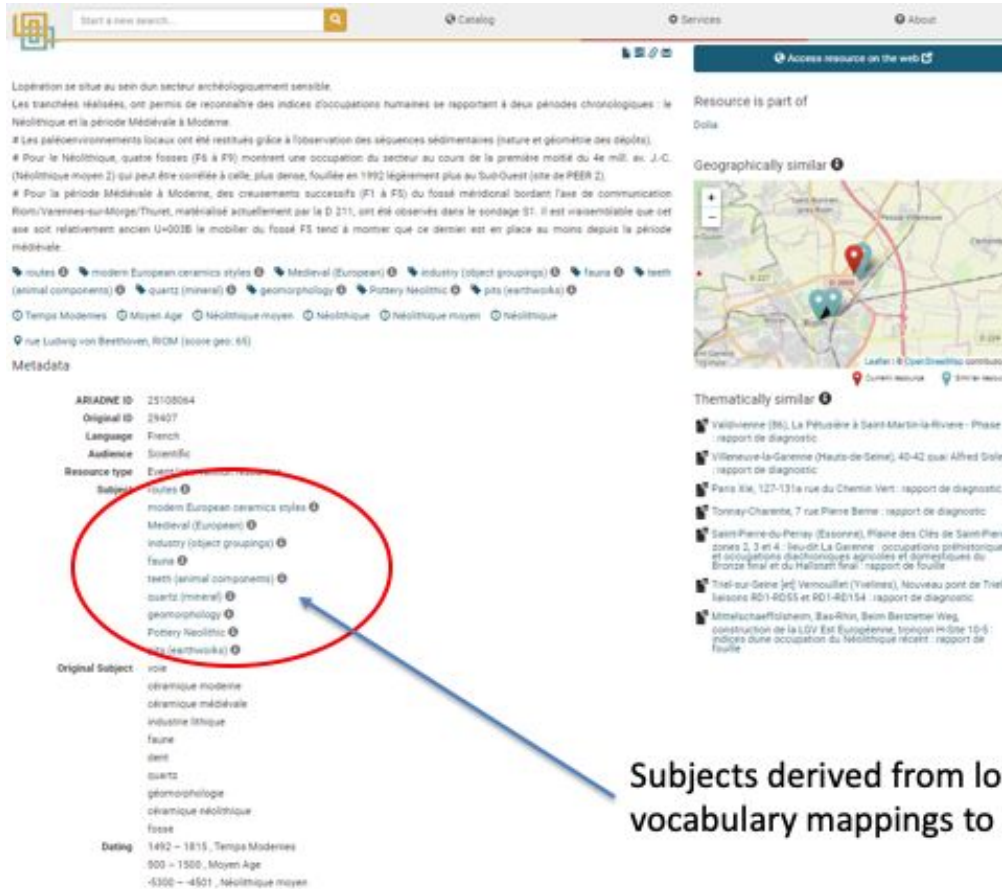
ARIADNE type-ahead suggestions



The screenshot shows the ARIADNE website interface. At the top, the ARIADNE logo is displayed. Below it, a search bar contains the text 'axe'. A dropdown menu shows suggestions for 'axe', including 'axes', 'axhammers', 'socketed axes', 'flanged axes', 'cleavers', 'ax heads', and 'axes'. Below the search bar, there is a 'Welcome' section with the text 'Explore the digital resources and learning and teaching.' and a 'Browse the Catalog' section. The 'Browse the Catalog' section includes a map of Europe, a bar chart showing the frequency of terms over time, and a list of terms including 'settlements', 'cist graves', 'buildings (structures)', 'barns', 'houses', 'hearth', 'pits (earthworks)', 'farms', and 'churches (buildings)'.

Getty AAT subject
term type-ahead
suggestions
during search

Records supplemented with AAT



Opération se situe au sein d'un secteur archéologiquement sensible.

Les tranchées réalisées, ont permis de reconnaître des indices d'occupations humaines se rapportant à deux périodes chronologiques : le Néolithique et la période Médiévale à Moderne.

■ Les paléoenvironnements locaux ont été restitués grâce à l'observation des séquences sédimentaires (nature et géométrie des dépôts).

■ Pour le Néolithique, quatre fosses (F6 à F9) montrent une occupation du secteur au cours de la première moitié du 4e mill. av. J.-C. (Néolithique moyen 2) qui peut être corrélée à celle, plus dense, fouillée en 1992 légèrement plus au Sud-Ouest (site de PEER 2).

■ Pour la période Médiévale à Moderne, des creusements successifs (F1 à F5) du fossé méridional bordent l'axe de communication Riom/Varennes-sur-Morge/Thuret, matérialisé actuellement par la D 211, ont été observés dans le sondage ST. Il est vraisemblable que cet axe soit relativement ancien U+001B le mobilier du fossé FS tend à montrer que ce dernier est en place au moins depuis la période médiévale.

Routes 1 modern European ceramics styles 1 Medieval (European) 1 Industry (object groupings) 1 faune 1 teeth (animal components) 1 quartz (minerals) 1 geomorphology 1 Pottery Neolithic 1 pits (earthworks) 1

Temps Modernes 1 Moyen Age 1 Néolithique-moyen 1 Néolithique 1 Néolithique-moyen 1 Néolithique

rue Luther von Beethoven, Riom (score geo: 65)

Metadata

ARIADNE ID: 2510064
Original ID: 25407
Language: French
Audience: Scientific
Resource type: Event

Subject 1
routes 1
modern European ceramics styles 1
Medieval (European) 1
Industry (object groupings) 1
faune 1
teeth (animal components) 1
quartz (minerals) 1
geomorphology 1
Pottery Neolithic 1
pits (earthworks) 1

Original Subject
voie
céramique moderne
céramique médiévale
industrie litique
faune
dent
quartz
géomorphologie
céramique néolithique
fosse

Dating
1492 - 1815, Temps Modernes
900 - 1500, Moyen Age
-5300 - -4501, Néolithique moyen

Resource is part of
Dolla

Geographically similar 1

Thematically similar 1

Vallée de la Seine (38), La Pélusie à Saint-Martin-la-Rivière - Phase 2 : rapport de diagnostic
Villeneuve-la-Garenne (Haut-de-Seine), 40-42 quai Alfred Sisley : rapport de diagnostic
Paris XIX, 127-131a rue du Chemin Vert : rapport de diagnostic
Tonnay-Charente, 7 rue Pierre Berner : rapport de diagnostic
Saint-Pierre-du-Perray (Esson), Plaine des Clés de Saint-Pierre, zones 2, 3 et 4 : lieu-dit La Garenne : occupations préhistoriques et occupations diagnostiques agricoles et symétriques du Bronze final et du Hallstatt final : rapport de fouille
Triel-sur-Seine (en Verneuilien (Yvelines), Nouveau pont de Triel : faience R01-R055 et R01-R0154 : rapport de diagnostic
Mittelbachweilshausen, Bas-Rhin, Beim Bernheimer Weg, construction de la LDV Est Européenne, tronçon M-Site 10-5 : indices d'une occupation du Néolithique récent : rapport de fouille

Subjects derived from local vocabulary mappings to AAT

Selected references and links

References

- Binding, C, Tudhope, D & Vlachidis, A 2018, 'A study of semantic integration across archaeological data and reports in different languages' Journal of Information Science, vol 45, no. 3, pp. 364-386. [doi:10.1177/0165551518789874](https://doi.org/10.1177/0165551518789874)
- Binding, C & Tudhope, D 2016, 'Improving interoperability using vocabulary linked data' International Journal on Digital Libraries, vol 17, no. 1, pp. 5-21. [doi:10.1007/s00799-015-0166-y](https://doi.org/10.1007/s00799-015-0166-y)

Links

- ARIADNEplus project: <http://www.ariadne-infrastructure.eu/>
- ARIADNE portal: <https://ariadne-infrastructure.eu/portal/>
- Vocabulary Matching Tool (VMT): <https://vmt.ariadne.d4science.org/vmt/>
- USW Hypermedia Research Group: <https://hypermedia.research.southwales.ac.uk/>

Contact

- ceri.binding@southwales.ac.uk ORCID: 0000-0002-6376-9613
- douglas.tudhope@southwales.ac.uk

Management of controlled vocabularies within the Aïoli platform with the Opentheso tool



WP4 - Innovations in Data Production

T4.6 Semantic annotation of Heritage Science Data

Isabelle Cao
CNRS-MAP

- Activities in WP4
- Aïoli platform
- Opentheso tool
- Implementation of Opentheso within the Aïoli platform
- Collected and integrated vocabularies
- Pursuits and prospects

Activities

WP4



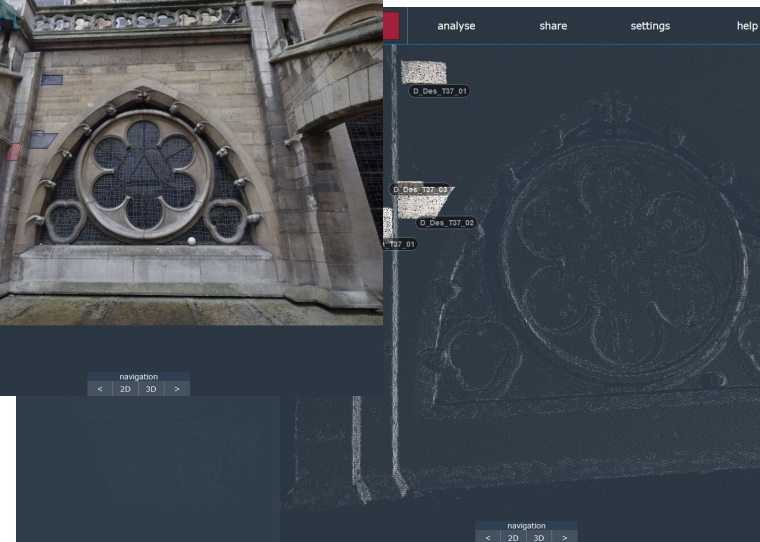
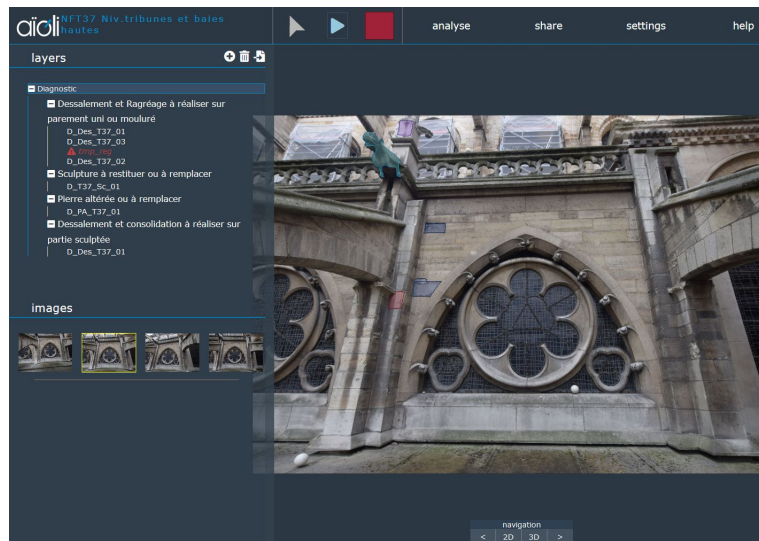
- D4.16 Specification of the new features of the Aioli platform
- Technical robustness of the platform
- Collaboration framework
- Management of controlled vocabularies
- Compatibility with the CIDOC-CRM (in collaboration with FORTH)

Aïoli

a reality-based 3D annotation platform



- Automatic Image-based 3D reconstruction
- 2D-3D Multi-layer annotation
- Custom user descriptors (dimensions, geometric and visual properties)
- Attachments to additional resources
- Online

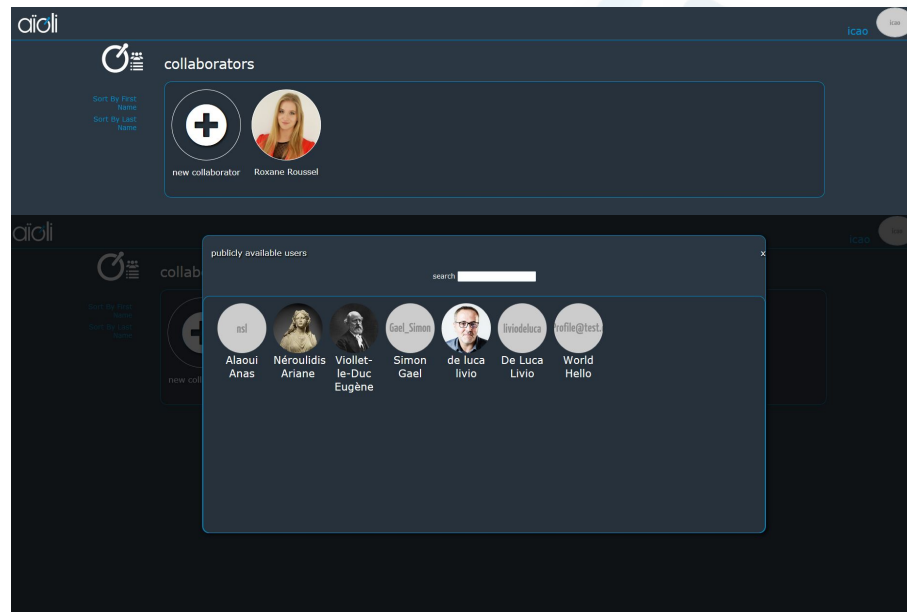


Aïoli

Integrating controlled vocabularies



- Multiple collaborators
- Consistency of terms
- W3C SKOS standard
- Link to Opentheso



Opentheso

a web-based thesaurus management tool

- MOM-CNRS (Maison de l'Orient et de la Méditerranée)
- Multilingual thesaurus management software
- ISO-25964 standard
- ongoing development
- opensource on Github
- Partners : Frantiq; MOM; MASA consortium; Huma-Num



Opentheso Features

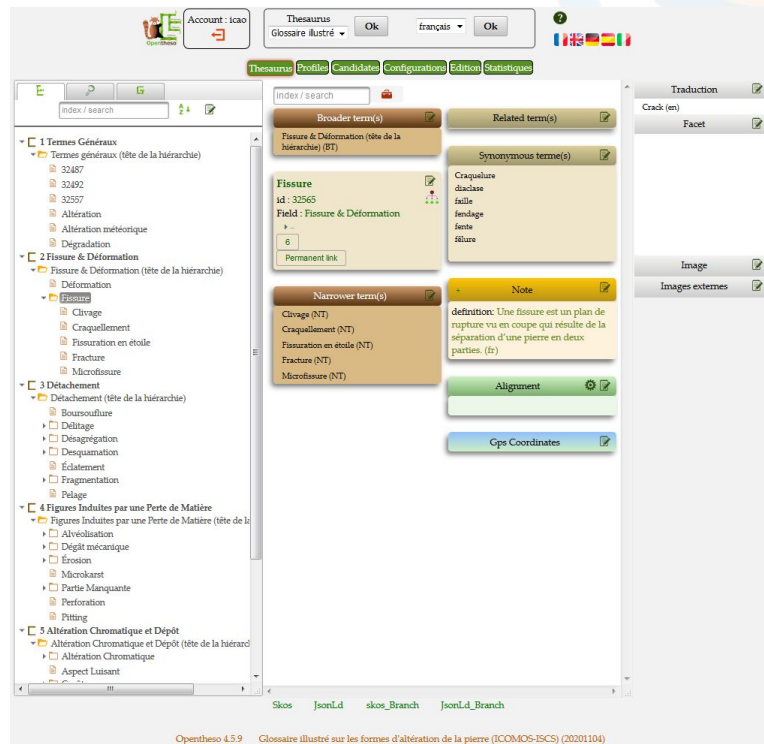
- Import/export RDF (SKOS, JsonLD, Turtle)
- Import CSV and list of terms
- Management of concepts
- Drag and Drop
- Alignment
- Collaborative
- API (REST WebServices)



Opentheso

Concepts management

- Add concepts
- Modify concepts
- Add synonyms
- Add notes
- Create hierarchical relationships
- Create associative relationships
- Translate concepts



The screenshot displays the Opentheso web interface for managing concepts. The interface is in French and shows a hierarchical tree of concepts on the left. The main panel displays details for the concept 'Fissure' (id: 32565) under the field 'Fissure & Déformation'. It includes sections for 'Broader term(s)', 'Related term(s)', 'Synonymous term(s)', and 'Narrower term(s)'. A 'Note' section provides a definition: 'définition: Une fissure est un plan de rupture vu en coupe qui résulte de la séparation d'une pierre en deux parties. (fr)'. The interface also features a 'Translation' panel on the right with options for 'Crack (en)', 'Facet', 'Image', and 'Images externes'. The bottom status bar indicates 'Opentheso 4.5.9' and 'Glossaire illustré sur les formes d'altération de la pierre (ICOMOS-ISCS) (20201104)'.

Opentheso

Thesaurus navigation

- Tree navigation
- Hierarchy view
- Collection view
- Index view



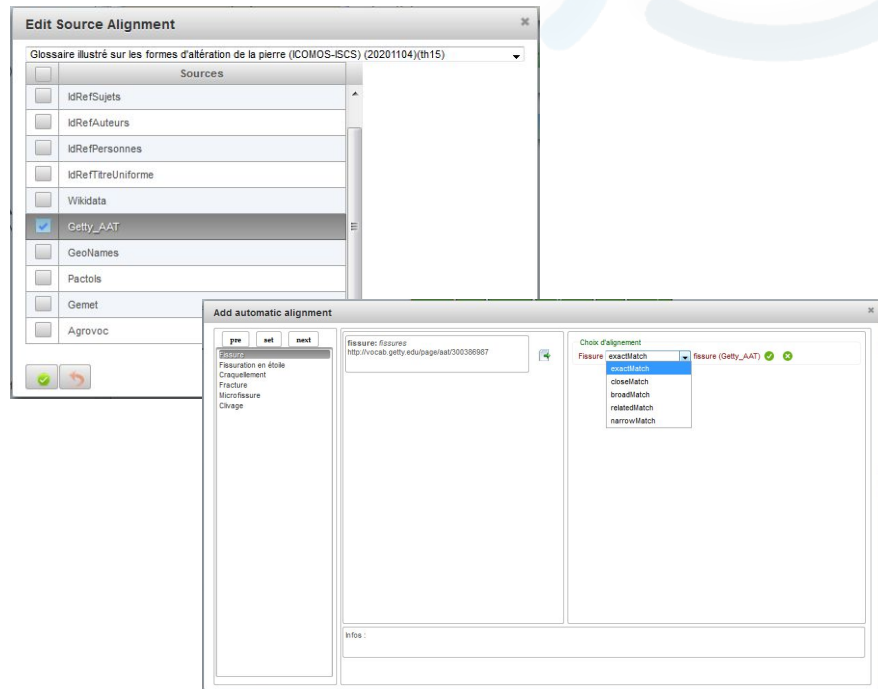
The screenshot displays the Opentheso interface with three main components:

- Tree navigation (left):** A hierarchical tree structure showing categories like '1 Termes Généraux', '2 Fissure & Déformation', '3 Détachement', '4 Figures Induites par une Perte de Matière', and '5 Altération Chromatique et Dépôt'.
- Hierarchy view (middle):** A list of terms under the selected category '2 Fissure & Déformation', including 'Fissuration en étoile', 'Craquellement', 'Fracture', 'Microfissure', 'Clivage', 'Fissuration en étoile', 'Déformation', 'Fissure & Déformation (tête de la hiérarchie)', and 'Fissure'.
- Search result table (bottom right):** A table showing search results for the term 'fissure'.

Name	Field
Fissure	Fissure & Déformation
Fissure & Déformation (tête de la hiérarchie)	Fissure & Déformation
Microfissure	Fissure & Déformation

Opentheso Alignment

- Semi-automatic alignment
- Several pre-configured sources
- Results return
- Different levels of equivalence



Opentheso

Collaborative features



- Candidate module
- Submit a concept
- Enrich with a note
- Enrich with a translation
- Automatic addition

Account : icao

Thesaurus Profiles Candidates Configurations Edition Statistiques

My candidates: tot = 1

Candidat edition

Currents	Insertion	Archives
	Nb	State
soulèvement(CA_1)	1	Pending

soulèvement nb : 1 State : Pending

ica0 Date: 2020-11-05

Note

Etat instable de la couche picturale se présentant sous forme d'écaillies d'une ou de plusieurs strates peu adhérentes soit entre elles, soit au support, souvent à cause d'une perte de cohésion d'une strate sous-jacente. (CMN)

Level (/)

add translation anglais (en) Ok

Language	Traduction	Author
No existing traduction.		

Opentheso 4.5.9 Glossaire illustré sur les formes d'altération de la pierre (ICOMOS-ISCS) (20201104)

Opentheso Collaborative features



- User role
- SuperAdmin, admin, manager and contributor
- Create projects
- Manage thesauri in projects
- Manage users in project

The screenshot displays the Opentheso web interface. At the top, there's a navigation bar with links: Thesaurus, Profiles, Candidates, Configurations, Edition, and Statistiques. The user is logged in as 'icaio' with the role 'superAdmin'. The main content area is divided into two sections: 'Users' and 'My account'.

Users Section:

Projets: Notre-Dame

Déplacer les thésos entre les groupes

liste thésaurus	liste groupes	
Architecture (Pérouse de Montclos)	Notre-Dame (1)	<input type="checkbox"/>
CICRP (th4)	LabCom CICRP-MAP (3)	<input type="checkbox"/>
Techniques d'analyses et d'exams	PARCOURS (4)	<input type="checkbox"/>
Thésaurus Notre-Dame de Paris (20)	Notre-Dame (1)	<input type="checkbox"/>
Vocabulaire 3D - Lexique pour les S	Consortium 3D SHS (2)	<input type="checkbox"/>
Glossaire illustré sur les formes d'alt		<input type="checkbox"/>
Notre-Dame références abrégées (20)		<input type="checkbox"/>
Thésaurus de la désignation des obj		<input type="checkbox"/>
Thésaurus de la désignation des oeu		<input type="checkbox"/>

My account Section:

icaio -> superAdmin

Modifier votre pseudo:

Current Password:

New Password:

Confirm password:

isabelle.cao@map.cnrs.fr

New address:

Alert by mail: ☒

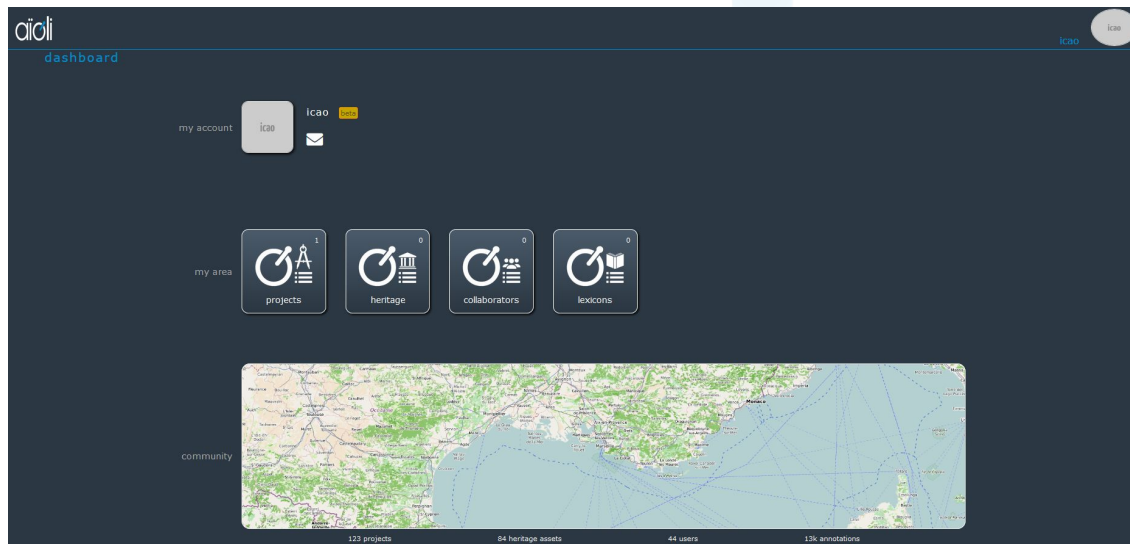
Users (Notre-Dame) Section:

Nickname	Rights				
francois.morlet	contributor	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
roxane.roussel	manager	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
violette.abergel	admin	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

Implementation

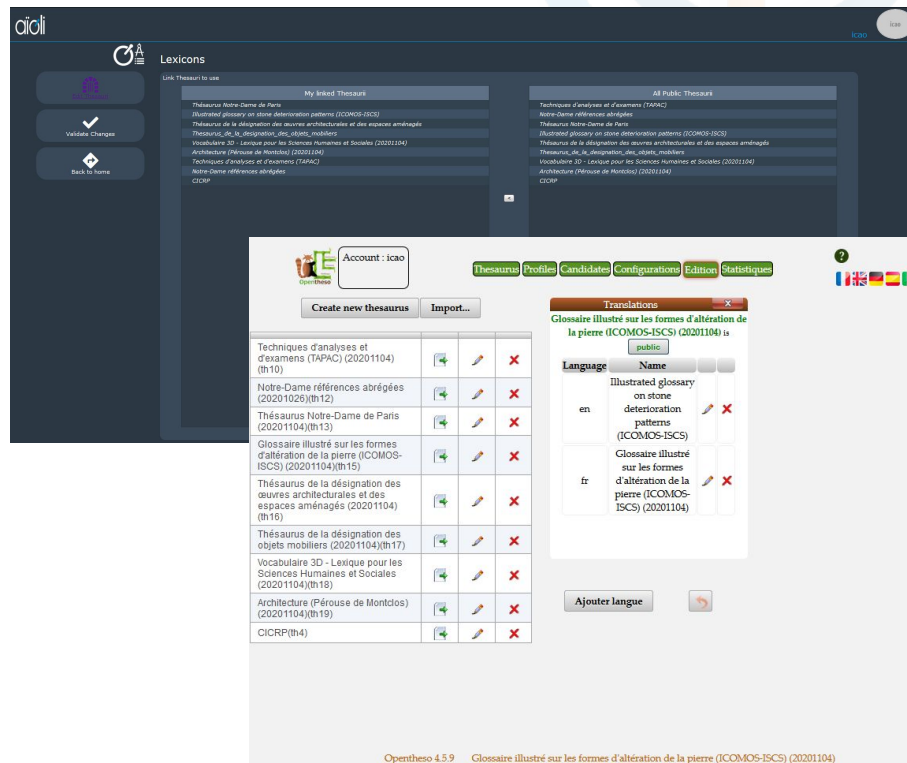
Technical aspects

- Docker container
- Interface elements
- Autocomplete



Implementation Lexicon

- Link to the Opentheso docker
- Automatic synchronization
- Thesauri public status
- Link thesauri to use



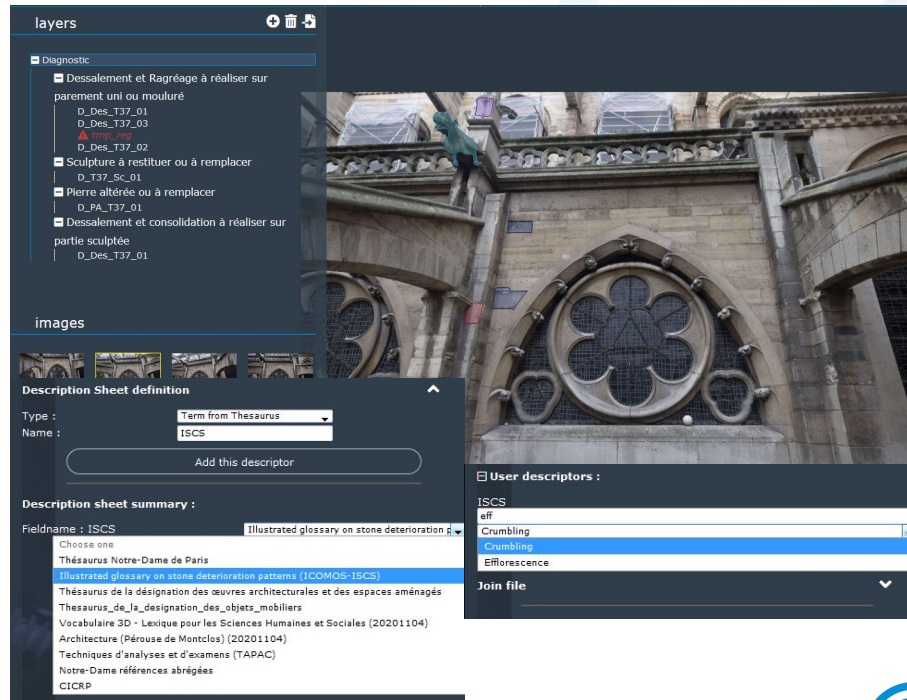
The screenshot displays the Opentheso web application. The top navigation bar includes the 'alio' logo, a search icon, and a 'Lexicons' dropdown menu. The main content area is divided into two panels: 'My linked Thesaurus' and 'All Public Thesaurus'. The 'My linked Thesaurus' panel lists several thesauri, including 'Thésaurus Notre-Dame de Paris', 'Illustrated glossary on stone deterioration patterns (ICOMOS-ISCS)', and 'Thésaurus de la désignation des œuvres architecturales et des espaces aménagés'. The 'All Public Thesaurus' panel lists similar thesauri, including 'Techniques d'analyses et d'examen (TAPAC)', 'Notre-Dame références abrégées', and 'Illustrated glossary on stone deterioration patterns (ICOMOS-ISCS)'. A modal window titled 'Translations' is open, showing a table of translations for the thesaurus 'Glossaire illustré sur les formes d'altération de la pierre (ICOMOS-ISCS) (20201104)'. The table has columns for 'Language' and 'Name'. The 'Language' column lists 'en' and 'fr'. The 'Name' column lists 'Illustrated glossary on stone deterioration patterns (ICOMOS-ISCS)' and 'Glossaire illustré sur les formes d'altération de la pierre (ICOMOS-ISCS) (20201104)'. The 'Translations' modal window also includes a 'public' checkbox and an 'Ajouter langue' button.

Language	Name
en	Illustrated glossary on stone deterioration patterns (ICOMOS-ISCS)
fr	Glossaire illustré sur les formes d'altération de la pierre (ICOMOS-ISCS) (20201104)

Implementation

Data container

- User descriptor field
- Select type “Term from Thesaurus”
- Select the thesaurus linked
- Type in at least 3 letters to search for a word



The screenshot displays a web-based data container interface. At the top, a 'layers' panel shows a 'Diagnostic' section with a tree structure of descriptors: 'Dessalement et Ragréage à réaliser sur parement uni ou mouluré' (with sub-descriptors D_Des_T37_01, D_Des_T37_03, D_Des_T37_02), 'Sculpture à restituer ou à remplacer' (with D_T37_Sc_01), 'Pierre altérée ou à remplacer' (with D_JA-T37_01), and 'Dessalement et consolidation à réaliser sur partie sculptée' (with D_Des_T37_01). Below this is an 'images' section with a grid of small image thumbnails. The main area features a large image of a stone structure with a circular window. Overlaid on this is a 'Description Sheet definition' form. The form has fields for 'Type' (set to 'Term from Thesaurus') and 'Name' (set to 'ISCS'), with an 'Add this descriptor' button. Below the form is a 'Description sheet summary' section with a 'Fieldname' field (set to 'ISCS') and a dropdown menu for 'Choose one' thesaurus. The dropdown list includes: 'Thésaurus Notre-Dame de Paris', 'Illustrated glossary on stone deterioration patterns (ICOMOS-ISCS)' (highlighted), 'Thésaurus de la désignation des œuvres architecturales et des espaces aménagés', 'Thésaurus de la désignation des objets mobiliers', 'Vocabulaire 3D - Lexique pour les Sciences Humaines et Sociales (20201104)', 'Architecture (Pérouse de Montclos) (20201104)', 'Techniques d'analyses et d'examen (TAPAC)', 'Notre-Dame références abrégées', and 'CICRP'. On the right side of the interface, a 'User descriptors' section shows a list with 'ISCS', 'eff', 'Crumbling', 'Crumbling' (highlighted), and 'Efflorescence'. At the bottom right is a 'Join file' button.

Vocabularies integrated CH domain



- Architecture (Pérouse de Montclos)
- Thésaurus de la désignation des œuvres architecturales et des espaces aménagés (*Thesaurus of the designation of architectural works and landscaped spaces*)
- Thésaurus de la désignation des objets mobiliers (*Thesaurus of the designation of movable objects*)
- Techniques d'analyses et d'examens (PARCOURS) (*Analysis and examination techniques*)
- Illustrated glossary on stone deterioration patterns (ICOMOS-ISCS)
- EwaGlos - European Illustrated Glossary of Conservation Terms for Wall Paintings and Architectural Surfaces
- Vocabulaire 3D - Lexique pour les Sciences Humaines et Sociales (Consortium 3D SHS) (*3D Vocabulary - Lexicon for the Humanities and Social Sciences*)

Pursuits and prospects



CLARIN

- Beta version
- Tests in use
- Fix of bugs
- Upgrade the docker
- Search for annotation within a project

- Thank you for listening !
- Any question ?

isabelle.cao@map.cnrs.fr



SSHOC

social sciences & humanities open cloud

BackBone Thesaurus, a model for thesauri interoperability

SSHOC Considerations for the Vocabulary Platforms

SSHOC Workshop

06 November 2020

Eleni Tsoulou(c)ha

Chryssoula Bekiari



SSHOC, "Social Sciences and Humanities Open Cloud", has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-04-2018, grant agreement # 823782.



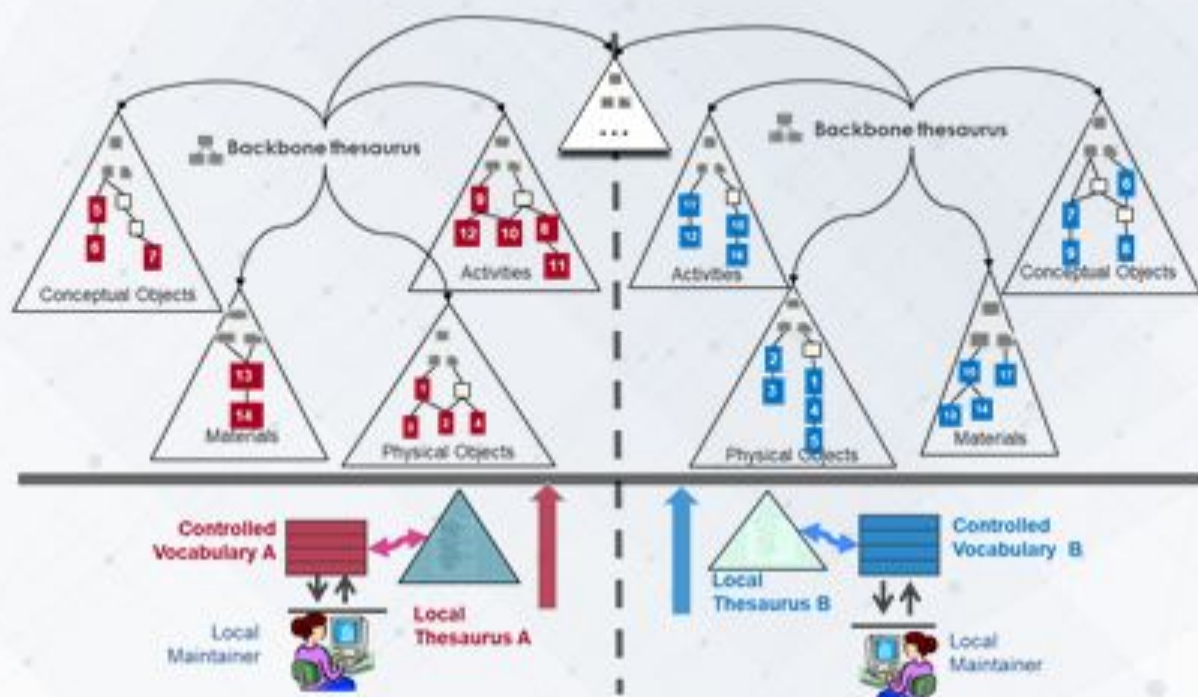
This project is funded from the EU Horizon 2020 Research and Innovation Programme (2014-2020) under Grant Agreement No. 823782

social sciences & humanities open cloud



Backbone Thesaurus - a Federation of thesauri

an effort of a coherent overarching thesaurus for the humanities, under which specialist thesauri and structured vocabularies used across scholarly communities can be aligned and form a thesaurus federation.



BackBone Thesaurus: A meta-thesaurus for the humanities

- ⊗ The BackBone Thesaurus (BBT) ^[1] was developed by the *DARIAH Thesaurus Maintenance WG* ^[2].
- ⊗ It is an overarching thesaurus:
 - ⊗ It supports a loose integration of multiple thesauri, by means of mapping them to a small set of high level terms (facets & hierarchies).
 - ⊗ It enables cross-disciplinary resource discovery and ensures compatibility with thesauri that cover highly specific scientific and developing areas of knowledge.
 - ⊗ Aligned vocabularies retain their autonomy and can be accessed through the global thesaurus.
- ⊗ The BBT relies on a **classification** considered **valid** and **consistent** from a cross-disciplinary perspective, inspired by the **high-level classes** of CIDOC-CRM.
- ⊗ The content of the federated thesauri can be accessed through the **ACDH vocabulary repository service** ^[3]

^[1] www.backbonethesaurus.eu

^[2] www.dariah.eu/activities/working_groups/thesaurus_maintenance

^[3] <https://vocabs.dariah.eu/en/>



The Structure of BBT

Activities (facet)

- activities (top term)
- - disciplines
- - functions
- - human interactions
- - intentional destructions

Conceptual Objects (facet)

- conceptual objects (top term)
- - concepts
- - methods
- - propositional objects
- - symbolic objects

Geometric Extents (facet)

- geometric extents (top term)
- - 3d-volumes
- - linear extents
- - points
- - surface areas

Geopolitical Units (facet)

- geopolitical units (top term)

Groups and Collectivities (facet)

- groups and collectivities (top term)

Material Things (facet)

- material things (top term)
- - built environment
- - mobile objects
- - physical features
- - structural parts of material objects

Materials (facet)

- materials (top term)

Natural Processes (facet)

- natural processes (top term)
- - geneses
- - natural disasters

Roles (facet)

- roles (top term)
- - offices
- - roles of interpersonal relations

Types of Epochs (facet)

- types of epochs (top term)





Methodological principles: The design of BBT

- BBT makes use of **very few** high-level concepts.
- BBT concepts were obtained through a **bottom-up analysis** of classificatory systems in use across the Humanities.
- Defining the BBT concepts relied on uncovering the **essential properties** of the concepts denoted by the terms of the vocabularies considered.
- BBT concepts comprise:
FACETS:
 mutually exclusive high-level concepts that appeal to general categories of human perception & are to some extent "objective" and valid from a cross disciplinary perspective.
HIERARCHIES:
 stand in a hierarchical *isA* relation to their respective top terms & can be differentiated by at least one salient feature.
- BBT assumes an **open-world position**: facets and hierarchies within them do not exhaustify the domain they classify.
- BBT is deliberately open to revision/expansion

IDN.welt Thesaurus » Aktivitäten

Disziplinen und Fachrichtungen

PREFERRED TERM

①

BROADER CONCEPT

Aktivitäten

NARROWER CONCEPTS

Afrikanische Archäologie
 Altamerikanische Archäologie
 Altamerikanistik
 Alte Geschichte
 Altertumswissenschaften
 Altorientalistik
 Analogiestudien
 Anthropologie
 Antikenrezeption
 Architekturgeschichte
 Architektur
 Archäoastronomie
 Archäobotanik
 Archäologie
 Archäometrie
 Archäozoologie

What do all the concepts listed under *Disziplinen und Fachrichtungen* have in common?

- ☐ they (roughly) correspond to professional/scientific branches
- ☐ they are exerted by individuals who have acquired the skills required to perform them
- ☐ they involve the adoption and application of (well-defined) domain-specific methods

How do the concepts listed under *Disziplinen und Fachrichtungen* differ from one another?

- ☐ they each refer to an area of professional/scientific expertise, well-demarcated by the subject it studies and the methods and practices it applies

Astrologie
 Astronomie
 Biblische Archäologie
 Bodenkunde
 Byzantinische Archäologie
 Example of bottom-up semantic decomposition (source: i-DAL.welt thesaurus)



BBTalk: A tool for managing the BBT

- ❖ BBTalk^[1] is an online service for managing thesauri integration with the BBT and updating its content, when necessary.
- ❖ BBTalk supports multilingual content: supported languages are **English, French, German and Greek**.
- ❖ Import/export in RDF
- ❖ BBTalk is developed and maintained by FORTH-ICS (www.ics.forth.gr), in the framework of the project "APOLLONIS Greek Infrastructure for Humanities, Arts and Language Research and Innovation (2017-2020)".

[1] <https://www.backbonethesaurus.eu/BBTalk/>



BBTalk: A tool for managing the BBT – thesaurus alignment

- BBTalk offers a thesaurus alignment functionality allowing end users:
 - declare relations between specialist thesauri and the BBT, via the Connections tab [Fig.].
 - provide evidence to confirm/disprove the high-level terms of the metathesaurus.

Form for creating a new connection of a local Thesaurus with the BBT

Submitter: **Popoulou** Submission Date: **04.11.2020** Version: **1.2.3**

BBT Term name*

Information related to connected term

Connected term name*

Connected term URI*

Connected term Relation

Information related to connected Thesaurus

Thesaurus Name

Thesaurus Submitter

Thesaurus Description

Thesaurus URI

Spang Endpoint

Thesaurus SKOS File URL

Connections SKOS File URL

New connection form

BBT version 1.2.3

Term Info

This term describes physical substances that are constituents of material objects or are used in their construction, but whose individual substance is not a factor in the objects' function (e.g. gold, water, bricks, etc.). The term may include pure raw materials, processed or unprocessed, and also materials that have been modified and are incorporated into objects. Their main attribute is that they cannot be individualised, that is, they cannot be divided 'under' with clear and distinct boundaries or rules in response to other units of the same or different kind.

Term Name Translation

matériaux

Translation of Source Word

Ce terme désigne les types de substances physiques qui composent les objets matériels ou sont utilisés dans leur fabrication mais dont la substance individuelle n'est pas un facteur déterminant de la fonction de l'objet (par exemple, le bois, l'acier etc.). Le terme peut inclure les matières premières, transformées ou non, mais aussi les matériaux qui ont été façonnés et intégrés dans des objets, sans être le principal et après ne pouvant pas être individualisés (c'est-à-dire qu'ils ne sont pas des unités aux frontières ou aux états rigides et distincts par rapport à des unités d'une nature physique ou chimique).

URI

https://orcid.org/parthenos_vocabularies/Concept/2678

Label

matériaux

Table of local Thesaurus Connections

Connected term	BBT term	Submitter	Submission Date	Connection Relation	Connection ID
Popoulou	matériaux	Popoulou	05.06.2019	Broader Match	2010
Materialien	matériaux	Lara	18.12.2018	Exact Match	2011
material element	matériaux	Popoulou	07.07.2018	Broader Match	2020
materially significant substance	matériaux	Popoulou	07.07.2018	Broader Match	2026
material substance	matériaux	Popoulou	07.07.2018	Broader Match	2027
mat	matériaux	Popoulou	07.07.2018	Broader Match	2028
materialia substance	matériaux	Popoulou	07.07.2018	Broader Match	2029

Links to BBT top-term "materials"



Terminologies aligned with BBT

(in part or fully):

- ❖ DYAS Humanities Thesaurus - <https://humanitiesthesaurus.academyofathens.gr>
- ❖ the iDAI.welt-thesaurus - <http://thesauri.dainst.org/en.htm>,
- ❖ PACTOLS - <https://pactols.frantiq.fr/opentheso>,
- ❖ the PARTHENOS Vocabularies - <https://isl.ics.forth.gr/bbt-federated-thesaurus/PARTHENOS/en>
- ❖ GEMET - <https://www.eionet.europa.eu/gemet/en/themes>
- ❖ Language of Bindings Thesaurus - <https://www.ligatus.org.uk/lob>.



<https://vocabs.dariah.eu/backbone-thesaurus/en/>

<https://pactohs.fandom.com/wiki/7dc=16083&id=TH>

Backbone Thesaurus

Alphabetical	Hierarchy
<p>1. <i>Introduction</i></p> <p>2. <i>Background</i></p> <p>3. <i>Methodology</i></p> <p>4. <i>Results</i></p> <p>5. <i>Discussion</i></p> <p>6. <i>Conclusion</i></p>	<p>1. <i>Introduction</i></p> <p>2. <i>Background</i></p> <p>3. <i>Methodology</i></p> <p>4. <i>Results</i></p> <p>5. <i>Discussion</i></p> <p>6. <i>Conclusion</i></p>

- 000001 activities
- 000010 disciplines
- 000001 human interactions
- 000002 intentional destructions
- 000001 functions
- 000002 natural processes
- 000003 materials
- 000004 material things
- 000004 geometric extents
- 000005 types of epochs
- 000006 conceptual objects
- 000007 groups and collectivities
- 000008 roles
- 000009 geopolitical units

The screenshot displays the OpenThesa application interface, which is a web-based thesaurus tool. The interface is organized into several key sections:

- Left Sidebar:** Contains navigation and metadata for the selected concept. It includes links to 'activities - disciplines', 'PROPOSED TERM', 'BROADER CONCEPT', 'SCOPE NOTE', 'RELATION TO GROUP', 'CREATOR', 'IN OTHER LANGUAGES', 'URI', and 'Download this concept'.
- Main Content Area:** Displays the '000010 disciplines' concept. It features a heading '000010 disciplines' and a detailed 'SCOPE NOTE' explaining the term's origin and usage. The note mentions that the term is used for 'activities or phenomena' and is related to 'philology'.
- Right Sidebar:** Contains a 'Concept' tab, a 'Group/collection' list, and a 'Total de la branche' section. The 'Concept' tab shows a 'Term: philologie' and a 'Group/collection: philologie (9)'. The 'Group/collection' list includes 'Concept générique', 'Concept spécifique', 'Concept associé', and 'Synonymes'. The 'Total de la branche' section shows a 'Definition' and a 'Traduction'.

An orange arrow points from the '000010 disciplines' heading in the main content area to the 'Concept' tab in the right sidebar, indicating the relationship between the concept and its classification.



Thank you for your attention

For more information (including documentation, tutorials), please refer to:

✿ <https://www.backbonethesaurus.eu/>

✿ <https://www.backbonethesaurus.eu/BBTalk/>

✿ <https://www.dariah.eu/activities/working-groups/thesaurus-maintenance/>





SSHOC

social sciences & humanities open cloud

Join our community



<https://www.sshopencloud.eu>



info@sshopencloud.eu



[@SSHOpenCloud](https://twitter.com/SSHOpenCloud)



[/in.sshopencloud](https://in.sshopencloud)



SSHOC, "Social Sciences and Humanities Open Cloud", has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-04-2018, grant agreement # 823782.

Controlled Vocabularies support in the Dataverse data repository

Slava Tykhonov

Data Archiving and Networked Services, the Netherlands

SSHOC webinar on Controlled Vocabularies, November 6, 2020

FAIR and Dataverse



FM [AID*]	Question	Dataverse Q'aire	Dataverse Optimized
Identifier type	1	DOI	DOI
F1A	2		
F1B	Not tested in Q'aire		
F2A	4A		
F2A	4B		
F3	5B		
F4	6A		
F4	6B		
A1.1	7A		
A1.2	8A		
A1.2	8B	N/A	N/A
A2	9		
I1	10		
I2	11		
I3	12		
R1.1	13		
R1.2	14A		

DATAVERSE FAIR SUMMARY

- Strong support for Findable, Accessible, and Reusable principles
- Weak for Interoperable principles
- In agreement* with FAIR test results (*F3 was fixed after test)
- There is no FAIR “compliance”
- Instead, it’s a process and can always be improved

Source:

[Mercè Crosas.](#)

[“FAIR principles and beyond: implementation in Dataverse”](#)

Out of the box CV support in Dataverse (1)

1	#metadataBlock	name	dataverseAlias	displayName	blockURI	
2		citation		Citation Metadata	https://dataverse.org	
3	#datasetField	name	title	description	watermark	fieldType
82	#controlledVocabula	DatasetField	Value	identifier	displayOrder	
83		subject	Agricultural Sciences	D01	0	
84		subject	Arts and Humanities	D0	1	
85		subject	Astronomy and Astr	D1	2	
86		subject	Business and Manag	D2	3	
87		subject	Chemistry	D3	4	
88		subject	Computer and Inform	D7	5	
89		subject	Earth and Environme	D4	6	
90		subject	Engineering	D5	7	
91		subject	Law	D8	8	
92		subject	Mathematical Scienc	D9	9	
93		subject	Medicine, Health and	D6	10	
94		subject	Physics	D10	11	
95		subject	Social Sciences	D11	12	
96		subject	Other	D12	13	
97		publicationIDType	ark		0	

Source: [Dataverse Metadata Schema](#)

Out of the box CV support in Dataverse (2)



Dataverse

Search ▾ User Guide Support Dataverse Admin **1** ▾

Subject

Keyword

Related Publication

Subject

Select...

- ☐ Agricultural Sciences
- ☐ Arts and Humanities
- ☐ Astronomy and Astrophysics
- ☐ Business and Management
- ☐ Chemistry
- ☐ Computer and Information Science
- ☐ Earth and Environmental Sciences
- ☐ Engineering

ID Type

Select...

ID Number

URL

Enter full URL, starting with http://

Notes

Depositor

Admin, Dataverse

Deposit Date

2020-11-05

Interoperability in EOSC

- **Technical interoperability** defined as the “ability of different information technology systems and software applications to communicate and exchange data”. It should allow “to accept data from each other and perform a given task in an appropriate and satisfactory manner *without the need for extra operator intervention*”.
- **Semantic interoperability** is “the ability of computer systems to transmit data with unambiguous, shared meaning. Semantic interoperability is a requirement to enable *machine computable logic, inferencing, knowledge discovery, and data*”.
- **Organisational interoperability** refers to the “way in which organisations align their business processes, responsibilities and expectations to achieve commonly agreed and mutually beneficial goals. Focus on the *requirements of the user community* by making services available, easily identifiable, accessible and user-focused”.
- **Legal interoperability** covers “the broader environment of *laws, policies, procedures* and cooperation agreements”

Source: [EOSC Interoperability Framework v1.0](#)

Global Research Identifier Database (GRID) in SKOS

```
<http://www.grid.ac/institutes/grid.1001.0> a skos:Concept ;
  rdfs:label "Australian National University"@en ;
  isni:id "0000 0001 2180 7477"@en ;
  dc:date "1946-01-01"@en ;
  dcterms:identifier "grid.1001.0"@en ;
  vivo:abbreviation "ANU"@en ;
  skos:code "grid.420434.5" ;
  skos:exactMatch "http://www.wikidata.org/entity/Q127990" ;
  skos:inScheme "http://www.grid.ac/schema#CS000" ;
  skos:memberOf cw:C0007 ;
  skos:prefLabel "Australian National University"@en ;
  vcard:Address "http://www.grid.ac/institutes/grid.1001.0/address-0" ;
  foaf:homepage "http://www.anu.edu.au/"@en .
```

```
<http://www.grid.ac/institutes/grid.1002.3> a skos:Concept ;
  rdfs:label "Monash University"@en ;
  isni:id "0000 0004 1936 7857"@en ;
  dc:date "1958-01-01"@en ;
  dcterms:identifier "grid.1002.3"@en ;
  skos:code "grid.420434.5" ;
  skos:exactMatch "http://www.wikidata.org/entity/Q598841" ;
  skos:inScheme "http://www.grid.ac/schema#CS000" ;
  skos:memberOf cw:C0007 ;
  skos:prefLabel "Monash University"@en ;
  vcard:Address "http://www.grid.ac/institutes/grid.1002.3/address-0" ;
  foaf:homepage "http://www.monash.edu/"@en .
```

```
<http://www.grid.ac/institutes/grid.10025.36> a skos:Concept ;
  rdfs:label "University of Liverpool"@en ;
  isni:id "0000 0004 1936 8470"@en ;
  dc:date "1882-01-01"@en ;
  dcterms:identifier "grid.10025.36"@en ;
  skos:code "grid.420434.5" ;
  skos:exactMatch "http://www.wikidata.org/entity/Q499510" ;
  skos:inScheme "http://www.grid.ac/schema#CS000" ;
  skos:memberOf cw:C0007 ;
  skos:prefLabel "University of Liverpool"@en ;
  vcard:Address "http://www.grid.ac/institutes/grid.10025.36/address-0" ;
  foaf:homepage "http://www.liv.ac.uk/"@en .
```

We already have a lot of data in the global Dataverse network.

Can we provide **depositors** a convenient web interface to link their metadata to external controlled vocabularies?

Is it possible to disambiguate concepts and create links automatically?

MeSH in SKOSMOS



finto Finnish Thesaurus and Ontology Service

Vocabularies About Feedback Help | suomi | svenska

Medical Subject Headings

Content language: English

Alphabetical Hierarchy

- Dicistroviridae
- Encephalitis Viruses
- Flaviviridae
- Flexiviridae
- Hepatitis Delta Virus
- Hepeviridae
- Leviviridae
- Luteoviridae
- Mononegavirales
- Nidovirales
 - Arteriviridae
 - Coronaviridae
 - Coronavirus
 - Coronavirus 229E, Human
 - Coronavirus NL63, Human
 - Coronavirus OC43, Human
 - Coronavirus, Bovine
 - Coronavirus, Canine
 - Coronavirus, Feline
 - Coronavirus, Rat
 - Coronavirus, Turkey
 - Infectious bronchitis virus
 - Middle East Respiratory Syndrome Coronavirus
 - Murine hepatitis virus
 - Porcine epidemic diarrhea virus
 - SARS Virus**
 - Transmissible gastroenteritis virus
- Roniviridae
- Nodaviridae
- Orthomyxoviridae
- Picobirnavirus
- Picornaviridae
- Potyvirus
- Reoviridae
- Retroviridae

... > RNA Viruses > Nidovirales > Coronaviridae > Coronavirus > SARS Virus

PREFERRED TERM **SARS Virus**

BROADER CONCEPT [Coronavirus](#)

ENTRY TERMS

- Coronavirus, SARS
- Coronavirus, SARS-Associated
- Coronavirus, SARS-Related
- Coronavirus, Urbani SARS-Associated
- SARS Associated Coronavirus
- SARS-Associated Coronavirus
- SARS-Associated Coronavirus, Urbani
- SARS Coronavirus
- SARS-CoV
- SARS Related Coronavirus
- SARS-Related Coronavirus
- Severe Acute Respiratory Syndrome Virus
- Urbani SARS Associated Coronavirus
- Urbani SARS-Associated Coronavirus

NOTE

SCOPE NOTE

HISTORY NOTE

IN OTHER LANGUAGES

URI

infection = SEVERE ACUTE RESPIRATORY SYNDROME

A species of CORONAVIRUS causing atypical respiratory disease (SEVERE ACUTE RESPIRATORY SYNDROME) in humans. The organism is believed to have first emerged in Guangdong Province, China, in 2002. The natural host is the Chinese horseshoe bat, RHINOLOPHUS sinicus.

2003

SARS-virus Finnish

SARS virus Swedish

<http://www.yso.fi/onto/mesh/D045473>

SKOS version of
Medical Subjects
Headings provided by
Finnish Thesaurus and
Ontology Service
(Finto):

<https://finto.fi/mesh/>

Multilingual support is
essential requirement!

SKOSMOS API specification in Swagger



Skosmos API

The Skosmos REST API is a read-only interface to the data stored on the vocabulary server. The URL namespace is the base URL of the Skosmos instance followed by `/rest/v1/`.

Most methods return the data as UTF-8 encoded JSON-LD, served using the `application/json` MIME type. The data consists of a single JSON object which includes JSON-LD context information (in the `@context` field) and one or more fields which contain the actual data. Some methods (`data`) return other formats (RDF/XML, Turtle, RDF/JSON) with the appropriate MIME type.

The API supports Cross-Origin Resource Sharing by setting the Access-Control-Allow-Origin HTTP header to `"*"` for all requests.

The API supports the JSONP convention of appending a callback parameter to any URL. The returned data will then be wrapped in a JavaScript function call using the function name provided as the callback parameter value. JSONP wrapped data will be served using the `application/javascript` MIME type.

Global methods

Show/Hide | List Operations | Expand Operations

Vocabulary-specific methods

Show/Hide | List Operations | Expand Operations

GET	/({vocid})/	General information about the vocabulary
GET	/({vocid})/types	Information about the types (classes) of objects in the vocabulary
GET	/({vocid})/topConcepts	Top concepts of the vocabulary
GET	/({vocid})/data	RDF data of the whole vocabulary or a specific concept. If the vocabulary has support for it, MARCXML data is available for the whole vocabulary in each language.
GET	/({vocid})/search	Finds concepts and collections from a vocabulary by query term
GET	/({vocid})/lookup	Look up concepts by label
GET	/({vocid})/vocabularyStatistics	Number of Concepts and Collections in the vocabulary
GET	/({vocid})/labelStatistics	Number of labels by language
GET	/({vocid})/index/	Initial letters of the alphabetical index

Source: [Finto API](#)

Skosmos has a quite long history and established ecosystem.

Skosmos API specification (protocol) is a candidate to be reused by other CV services for the technical interoperability:

- CV Service (CESSDA)
- Semantic Gateway (DANS)
- Network Digitaal Erfgoed (NDE)
- ...

Dataverse should support the same protocol!

Semantic Gateway (in development)



Dataverse CVM Setting Generator

Name

URL TSV

Submit

Upload your metadata block tsv file:

Choose File no file selected

Upload!

Organisation

<input type="checkbox"/> cessda	<input type="checkbox"/> thesaurus	<input type="checkbox"/> unesco	<input checked="" type="checkbox"/> grid
<input type="checkbox"/> mesh	<input type="checkbox"/> iptc	<input type="checkbox"/> agrovoc	<input type="checkbox"/> faechersystematik

InterviewKeyWords

<input type="checkbox"/> cessda	<input checked="" type="checkbox"/> thesaurus	<input checked="" type="checkbox"/> unesco	<input type="checkbox"/> grid
<input type="checkbox"/> mesh	<input type="checkbox"/> iptc	<input type="checkbox"/> agrovoc	<input type="checkbox"/> faechersystematik

Gateway URL

download Dataverse URL unblock-key **push**

Source: [Dataverse gateway](#)

Semantic Gateway configuration



main semantic-gateway / conf / gateway.xml Go to file

4tikhonov Configuration updated Latest commit 19e3440 7 days ago History

1 contributor

102 lines (102 sloc) 3.34 KB Raw Blame

```
<?xml version="1.0" encoding="UTF-8"?>
<vocabularies>
  <ontology name="cessda">
    <type>cessda</type>
    <api>https://vocabularies.cessda.eu</api>
    <uri>v1/suggest/Vocabulary/$vocab/version/1.0/language/en/limit/10/query/$term</uri>
  </ontology>
  <ontology name="Unit of Analysis">
    <type>cessda</type>
    <vocabulary>Unit of Analysis</vocabulary>
    <api>https://vocabularies.cessda.eu</api>
    <uri>v1/suggest/Vocabulary/$vocab/version/1.0/language/en/limit/10/query/$term</uri>
  </ontology>
  <ontology name="thesaurus">
    <type>skosmos</type>
    <vocabulary>thesaurus</vocabulary>
    <api>https://vocabulaires.irstea.fr</api>
    <uri>skosmos/rest/v1/search</uri>
    <parameters>
      <vocab>$vocab</vocab>
      <query>$term</query>
      <lang>$lang</lang>
    </parameters>
  </ontology>
  <ontology name="wikidata">
    <type>nde</type>
    <vocabulary>wikidata</vocabulary>
    <api>http://demo.netwerkdigitaalrfgood.nl:8080</api>
    <uri>nde/graphql</uri>
    <query>query=%20%7B%20terms(match%3A%22$term%22%2Cdataset%3A%5B%22$vocab%22%5D)%20%7B%20dataset%20terms%20%7Buri%2C%20altLabel%7D%20%7D%20%7D</query>
    <parameters>
      <vocab>$vocab</vocab>
      <query>$term</query>
    </parameters>
  </ontology>

```

Dataverse deposit form with ontologies



Dataverse Search User Guide Support Dataverse Admin

Geographic Coverage

Country / Nation Amsterdam State/Province +

City Other +

Geographic Unit +

Geographic Bounding Box

West Longitude East Longitude +

North Latitude South Latitude +

Unit of Analysis

Vocabulary Unit of Analysis Term +

unl
thesaurus
grid
agrovoc

Universe

Time Method

Vocabulary Time Method Term +

unesco fam

VocabularyURL

Every field can be linked to the appropriate controlled vocabularies in FAIR way!

The same metadata field linked to many ontologies



Dataverse Search User Guide Support **Dataverse Admin**

Subject * ?

Keyword ?

Vocabulary ? <input type="text" value="unesco"/>	Term ? <input type="text" value="Family"/>	<input type="button" value="+"/> <input type="button" value="-"/>
Vocabulary URL ? <input type="text" value="http://skos.um.es/unescothes/C01489"/>		
Vocabulary ? <input type="text" value="thesaurus"/>	Term ? <input type="text" value="family labour"/>	<input type="button" value="+"/> <input type="button" value="-"/>
Vocabulary URL ? <input type="text" value="http://vocabulaires.irstea.fr/thesaurus/T"/>		
Vocabulary ? <input type="text" value="agrovoc"/>	Term ? <input type="text" value="families"/>	<input type="button" value="+"/> <input type="button" value="-"/>
Vocabulary URL ? <input type="text" value="http://aims.fao.org/aos/agrovoc/c_2785"/>		
Vocabulary ? <input type="text" value="iptc"/>	Term ? <input type="text" value="fam"/>	<input type="button" value="+"/> <input type="button" value="-"/>
Vocabulary URL ? <input type="text"/>	<ul style="list-style-type: none">• family• family planning• famine	

Language switch will change the language of suggested terms!

Semantic Gateway lookup API

Scenario: when user selects vocabulary and search for some term, API will get filled values and return back the list of concepts in the Skosmos format:

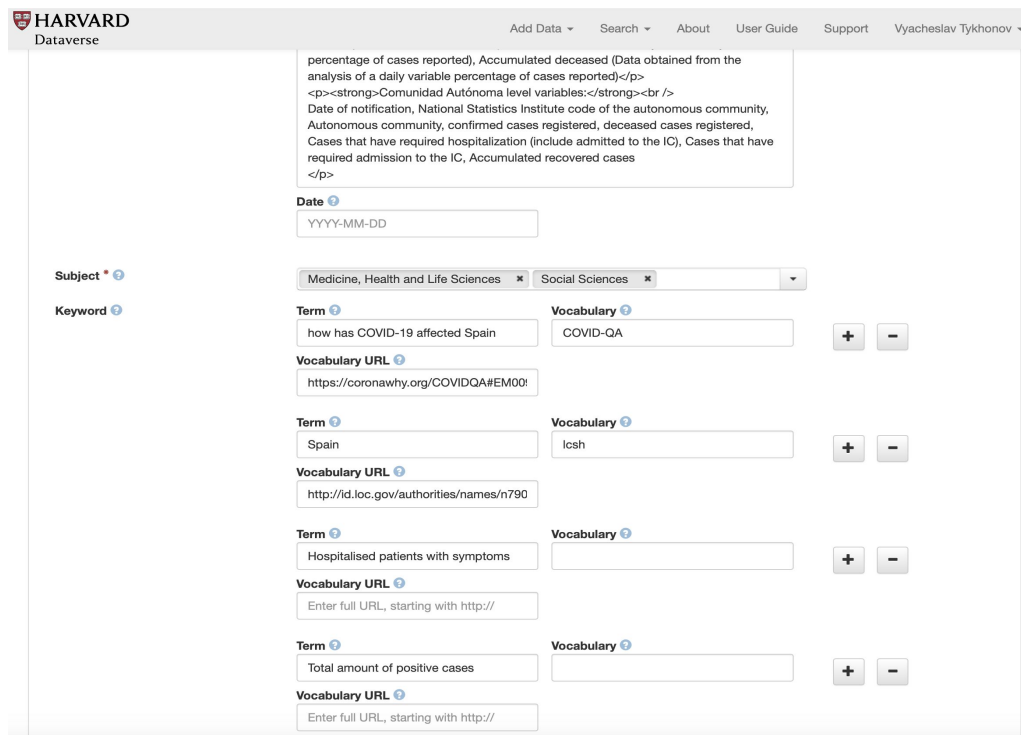
GET `/?lang=$language&vocab=$vocabulary&query=$keyword`

example:

GET [/?lang=en&vocab=unesco&query=fam](#)

Dataverse can be connected to any service with Skosmos protocol!

COVID-19 questions in Dataverse metadata



The screenshot shows the Harvard Dataverse interface. At the top, there's a navigation bar with links: Add Data, Search, About, User Guide, Support, and Vyacheslav Tykhonov. The main content area displays a dataset titled "percentage of cases reported), Accumulated deceased (Data obtained from the analysis of a daily variable percentage of cases reported)". Below the title, there's a "Date" field with a placeholder "YYYY-MM-DD". To the left, there are "Subject" and "Keyword" filters. The "Subject" filter is set to "Medicine, Health and Life Sciences" and "Social Sciences". The "Keyword" filter is empty. Below the subject filters, there are several "Term" and "Vocabulary" fields. The "Term" fields contain "how has COVID-19 affected Spain", "Spain", "Hospitalised patients with symptoms", and "Total amount of positive cases". The "Vocabulary" fields contain "COVID-QA", "lcsb", and are empty. Each "Term" field has a corresponding "Vocabulary URL" field. The "Vocabulary URL" fields contain "https://coronawhy.org/COVIDQA#EM00I", "http://id.loc.gov/authorities/names/n790", and are empty. There are also "+" and "-" buttons next to each "Vocabulary" field.

- COVID-19 ontologies can be hosted by SKOSMOS framework
- Researchers can enrich metadata by adding standardized questions provided by SKOSMOS ontologies
- Rich metadata exported back to Linked Open Data Cloud to increase a chance to be found
- Enriched metadata can be used for further ML models training



Source: [COVID-19 European data hub in Harvard Dataverse](#)

Thank you! Questions?



Slava Tykhonov
Senior Information Scientist
DANS-KNAW, the Netherlands
vyacheslav.tykhonov@dans.knaw.nl

Join our community



www.sshopencloud.eu



[@SSHOpenCloud](https://twitter.com/SSHOpenCloud)



[/in.sshopencloud](https://in.sshopencloud)



info@sshopencloud.eu

