ÖAW AUSTRIAN ACADEMY OF SCIENCES

austrian centre for digital humanities & cultural heritage

Vocabularies in the SSH Open Marketplace

TRIPLE training session, Berlin, 2023-03-27

Klaus Illmayer (OEAW ACDH-CH, klaus.illmayer@oeaw.ac.at)

ToC

- Overview on SSH Open Marketplace (SSHOMP) vocabularies
- Technical integration of vocabularies into SSHOMP
- Development of SSHOMP vocabularies
- User experience
- Curation of vocabularies
- Analysis of usage
- Experiences/Problems/Challenges
- Take aways

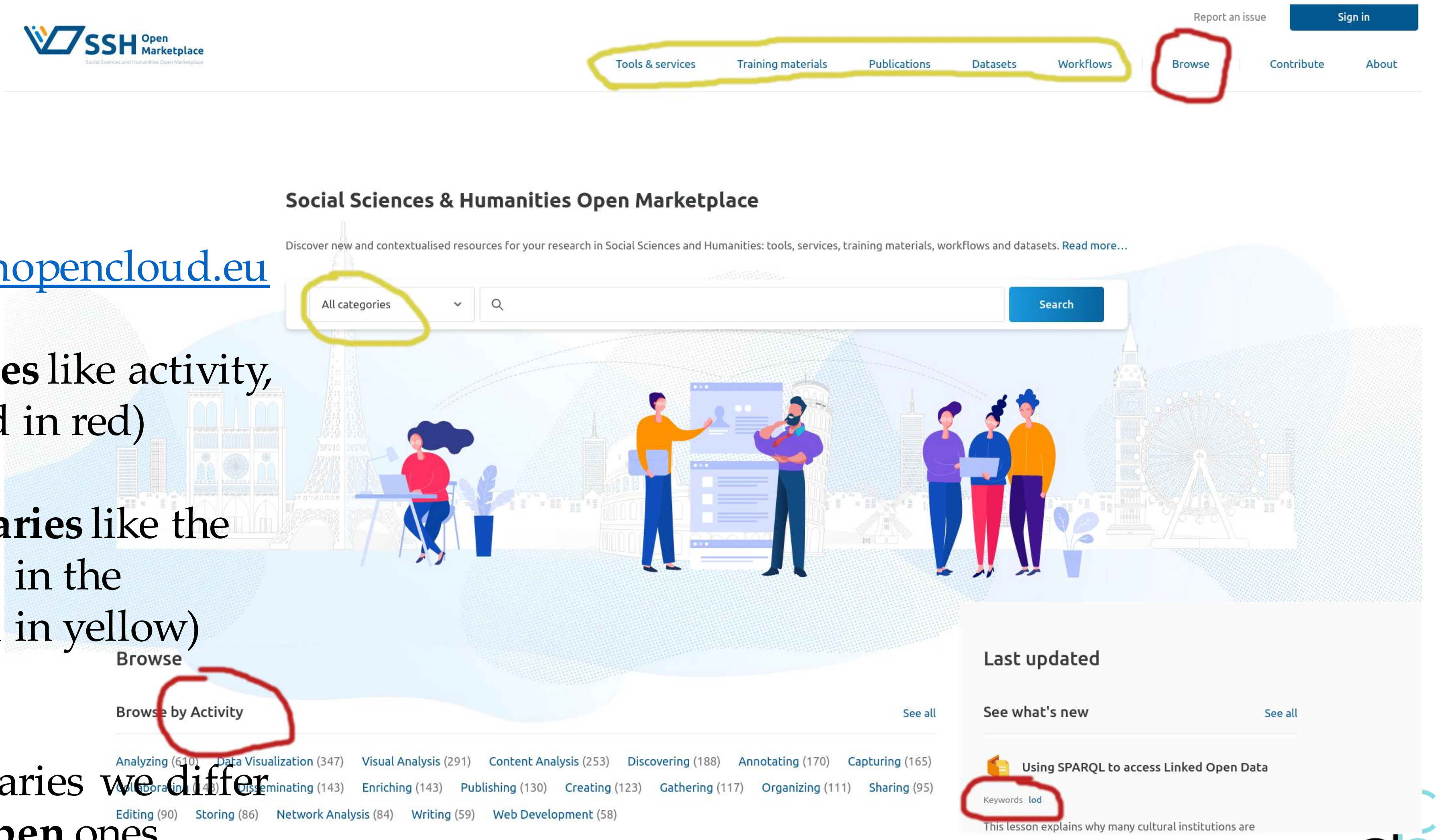
Overview

https://marketplace.sshopencloud.eu

Dedicated vocabularies like activity, keywords (highlighted in red)

but also **tacit vocabularies** like the categories represented in the SSHOMP (highlighted in yellow)

For dedicated vocabularies we differ between **closed** and **open** ones



SSH Open Marketplace

Report an issue | Sign in

Tools & services | Training materials | Publications | Datasets | Workflows | Browse | Contribute | About

Social Sciences & Humanities Open Marketplace

Discover new and contextualised resources for your research in Social Sciences and Humanities: tools, services, training materials, workflows and datasets. Read more...

All categories | Search

Browse

Browse by Activity — See all

Analyzing (610) Data Visualization (347) Visual Analysis (291) Content Analysis (253) Discovering (188) Annotating (170) Capturing (165) Collaborating (148) Disseminating (143) Enriching (143) Publishing (130) Creating (123) Gathering (117) Organizing (111) Sharing (95) Editing (90) Storing (86) Network Analysis (84) Writing (59) Web Development (58)

Last updated

See what's new — See all

Using SPARQL to access Linked Open Data

Keywords lod

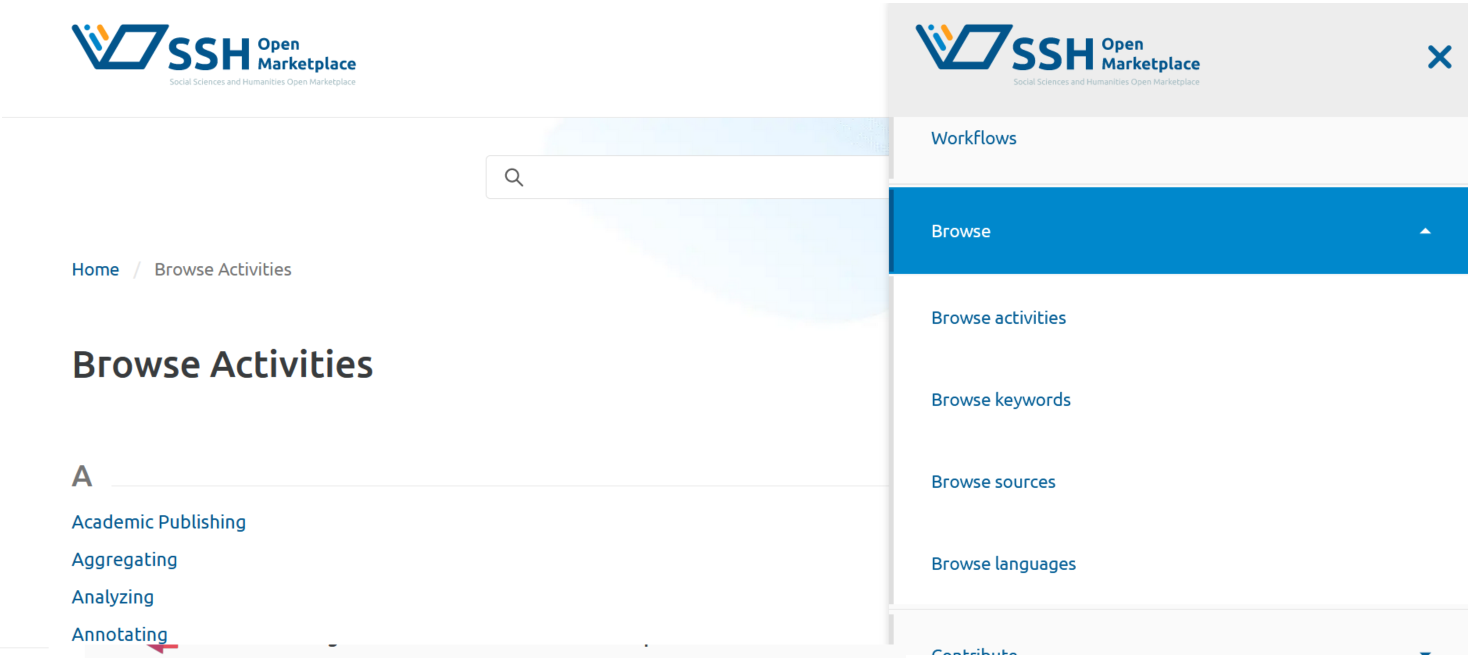
This lesson explains why many cultural institutions are

ÖAW AUSTRIAN ACADEMY OF SCIENCES

Overview

Using (some) vocabularies for **browsing**

… and for **facets**

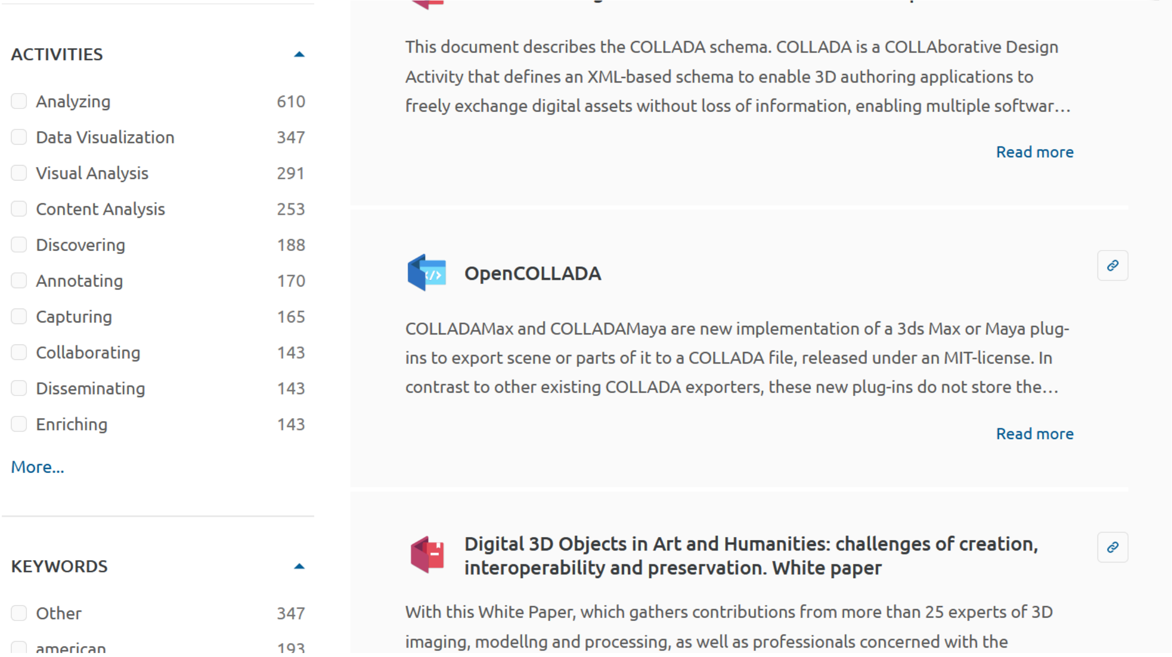


Home / Browse Activities

Browse Activities

A

Academic Publishing
Aggregating
Analyzing
Annotating



□ Analyzing	610
□ Data Visualization	347
□ Visual Analysis	291
□ Content Analysis	253
□ Discovering	188
□ Annotating	170
□ Capturing	165
□ Collaborating	143
□ Disseminating	143
□ Enriching	143

More...

□ Other	347
□ american	193

This document describes the COLLADA schema. COLLADA is a COLLAborative Design Activity that defines an XML-based schema to enable 3D authoring applications to freely exchange digital assets without loss of information, enabling multiple softwar…

Read more

OpenCOLLADA

COLLADAMax and COLLADAMaya are new implementation of a 3ds Max or Maya plug-ins to export scene or parts of it to a COLLADA file, released under an MIT-license. In contrast to other existing COLLADA exporters, these new plug-ins do not store the…

Read more

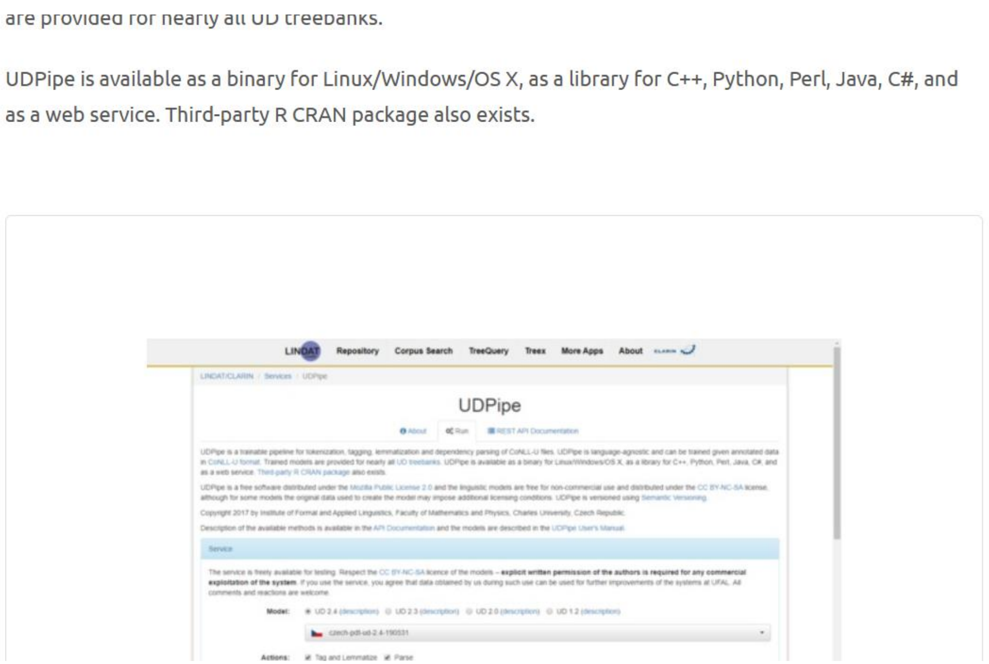
Digital 3D Objects in Art and Humanities: challenges of creation, interoperability and preservation. White paper

With this White Paper, which gathers contributions from more than 25 experts of 3D imaging, modellng and processing, as well as professionals concerned with the

Overview

Using vocabularies to **express detail information/metadata about an item**

… entered in a **structured format** based on the defined data schema for SSHOMP items



are provided for nearly all UD treebanks.

UDPipe is available as a binary for Linux/Windows/OS X, as a library for C++, Python, Perl, Java, C#, and as a web service. Third-party R CRAN package also exists.

LINDAT Repository Corpus Search TreeQuery Treex More Apps About

LINDAT/CLARIN / Services / UDPipe

UDPipe

About Run REST API Documentation

UDPipe is a trainable pipeline for tokenization, tagging, lemmatization and dependency parsing of CoNLL-U files. UDPipe is language-agnostic and can be trained given annotated data in CoNLL-U format. Trained models are provided for nearly all UD treebanks. UDPipe is available as a binary for Linux/Windows/OS X, as a library for C++, Python, Perl, Java, C#, and as a web service. Third-party R CRAN package also exists.

UDPipe is a free software distributed under the Mozilla Public License 2.0 and the linguistic models are free for non-commercial use and distributed under the CC BY-NC-SA license, although for some models the original data used to create the model may impose additional licensing conditions. UDPipe is versioned using Semantic Versioning.

Copyright 2017 by Institute of Formal and Applied Linguistics, Faculty of Mathematics and Physics, Charles University, Czech Republic.

Description of the available methods is available in the API Documentation and the models are described in the UDPipe User's Manual.

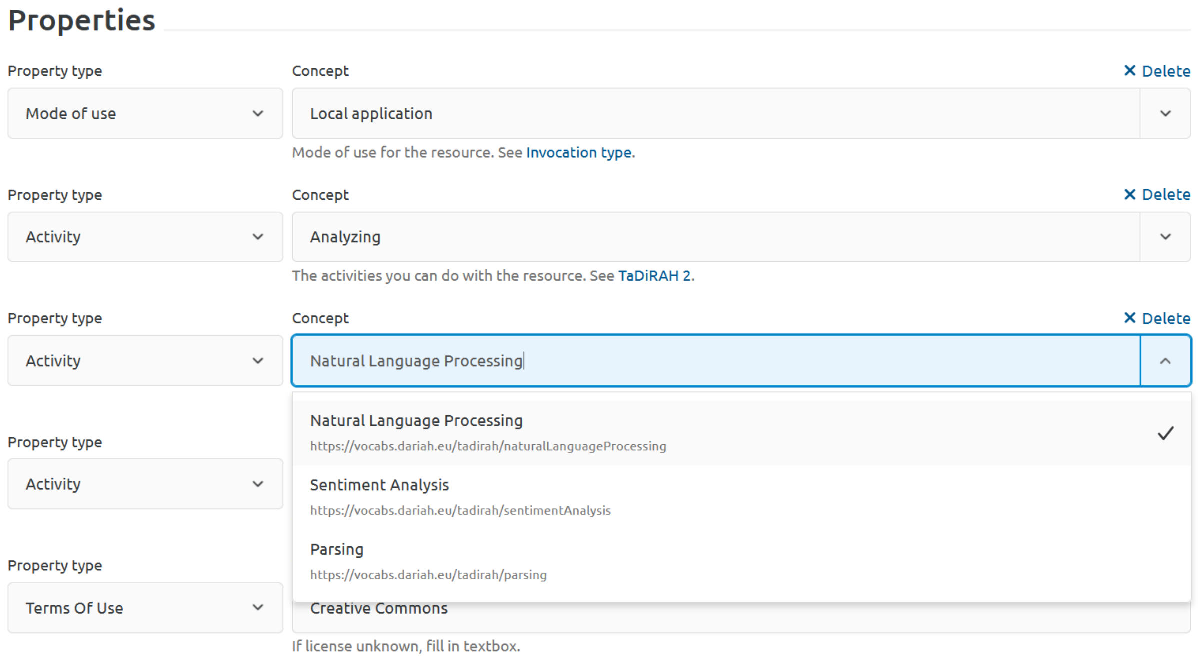
Service

The service is freely available for testing. Respect the CC BY-NC-SA licence of the models – **explicit written permission of the authors is required for any commercial exploitation of the system**. If you use the service, you agree that data obtained by us during such use can be used for further improvements of the systems at UFAL. All comments and reactions are welcome.

Model: ● UD 2.4 (description) ○ UD 2.3 (description) ○ UD 2.0 (description) ○ UD 1.2 (description)

czech-pdt-ud-2.4-190531

Actions: ☑ Tag and Lemmatize ☑ Parse



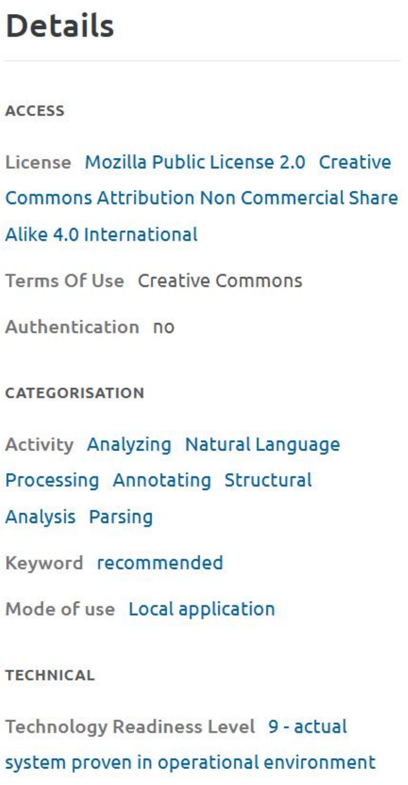
Properties

Property type	Concept
Mode of use	Local application Mode of use for the resource. See Invocation type.
Activity	Analyzing The activities you can do with the resource. See TaDiRAH 2.
Activity	Natural Language Processing
Activity	
Terms Of Use	Creative Commons If license unknown, fill in textbox.

Natural Language Processing ✓
https://vocabs.dariah.eu/tadirah/naturalLanguageProcessing

Sentiment Analysis
https://vocabs.dariah.eu/tadirah/sentimentAnalysis

Parsing
https://vocabs.dariah.eu/tadirah/parsing



Details

ACCESS

License Mozilla Public License 2.0 Creative Commons Attribution Non Commercial Share Alike 4.0 International

Terms Of Use Creative Commons

Authentication no

CATEGORISATION

Activity Analyzing Natural Language Processing Annotating Structural Analysis Parsing

Keyword recommended

Mode of use Local application

TECHNICAL

Technology Readiness Level 9 - actual system proven in operational environment

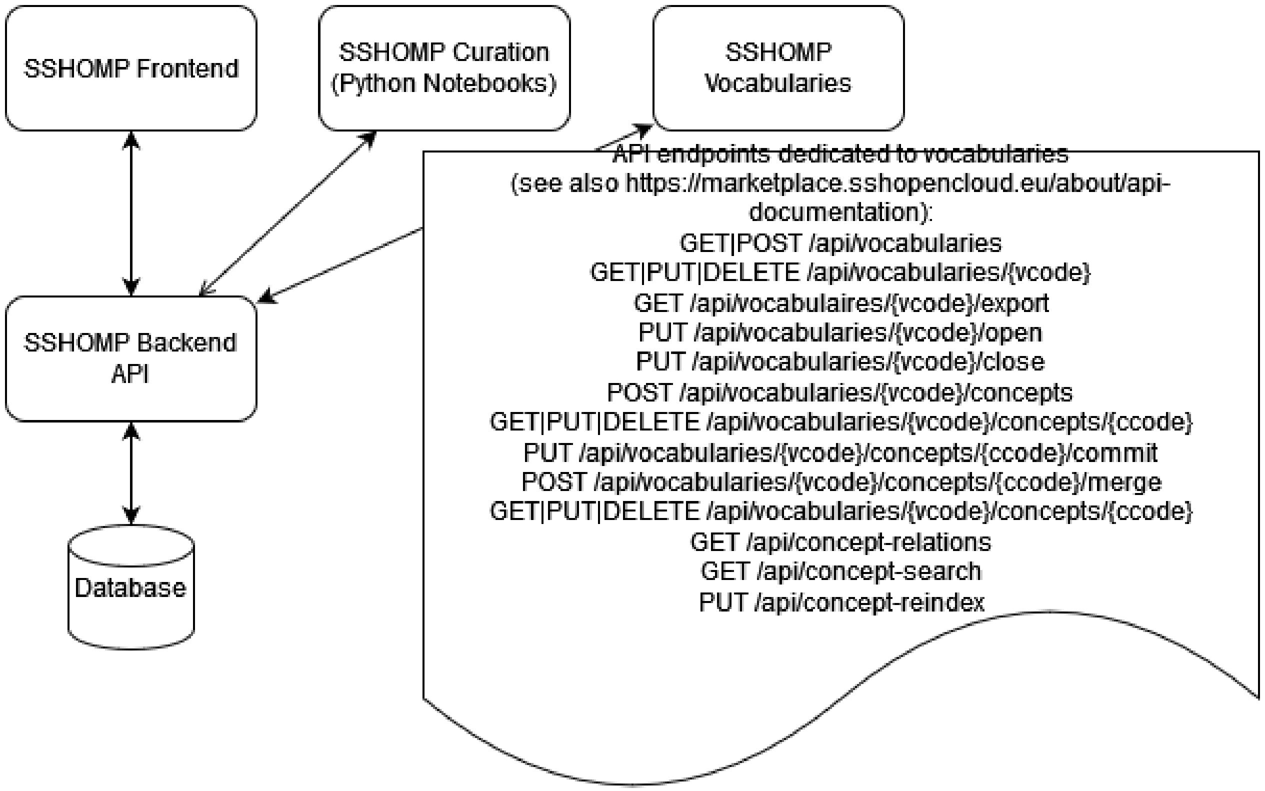
Overview

- Vocabularies expressed in SKOS are **supportive** for sharing information on items in the SSHOMP
- **Users** are asked to **add metadata on items** (= focus of SSHOMP), some of these fields (= properties we call them in SSHOMP) are vocabularies (= side effect of SSHOMP) > vocabulary/concepts should give **context** but they itself don't have much context
- Sometimes we do have a **field based on a vocabulary next to a text field** to use if no proper concept was found, e.g., for license we use this approach
- **Closed** (= controlled) vocabularies are **preferred** but for **user interaction** there is **one open vocabulary** ("sshoc-keyword") where users **tag** items and **propose new concepts**
- UX nudges **users to use fields** based on controlled vocabularies
- **Curation** and **ingestion** are brokers between open & closed vocabularies: ingestion **maps** to vocabularies, curation manually **post-processes** better usage of vocabularies
- No machine learning behind curation > **simple but pragmatic approaches** instead
- SSHOMP does not have a **curated keyword vocabulary** (like conversion-hub/TDT have)

Technical integration

Architecture:

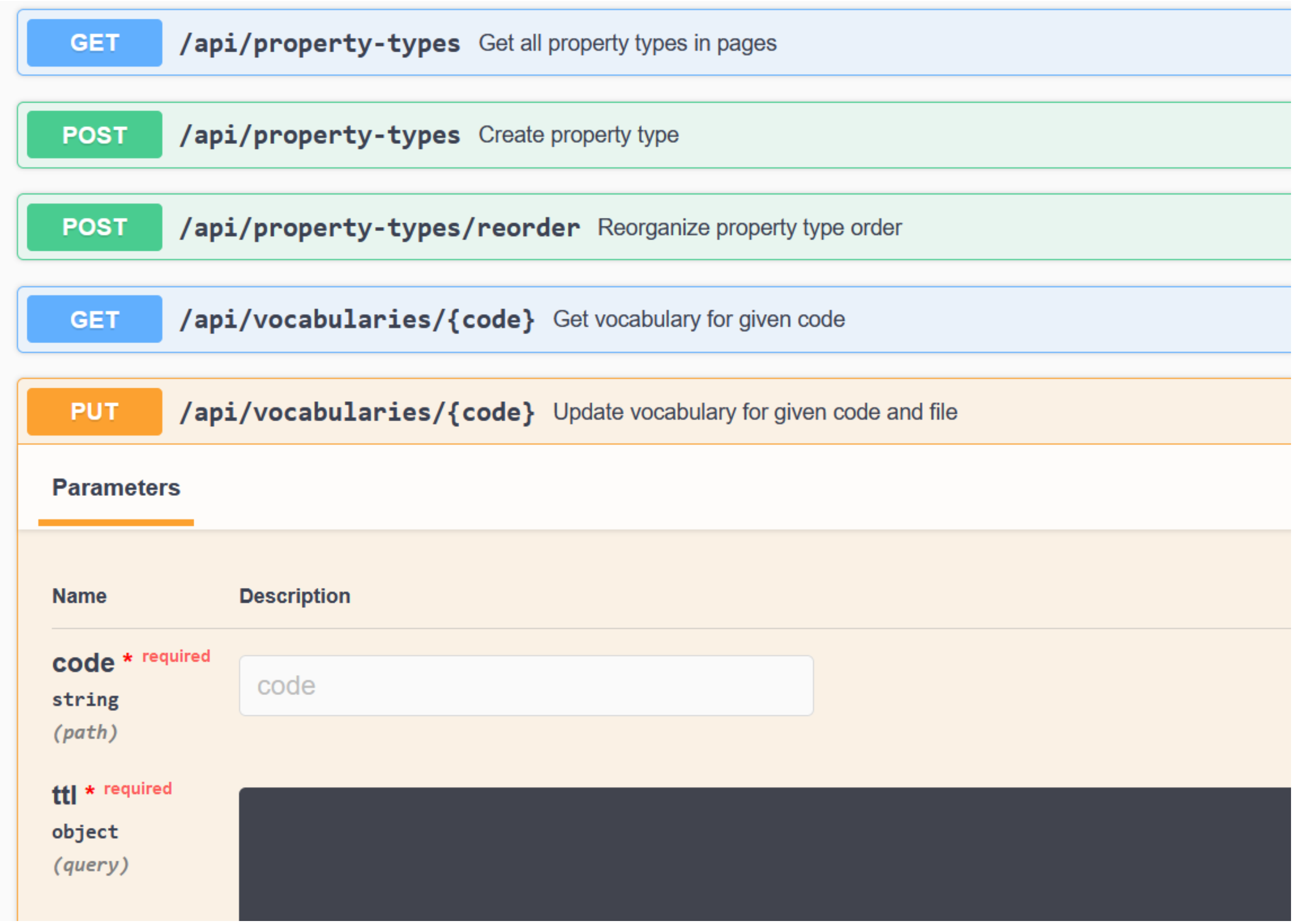
- **Decoupled** approach
- Vocabularies in **database**
- … and indexed in **Solr**
- Dedicated **API endpoints** (REST) for vocabularies
- ...that we use for **curation**
- ...but can be also used by others: **read access** is open and free, **write operations** need an authorization



API endpoints dedicated to vocabularies (see also https://marketplace.sshopencloud.eu/about/api-documentation):
GET|POST /api/vocabularies
GET|PUT|DELETE /api/vocabularies/{vcode}
GET /api/vocabulaires/{vcode}/export
PUT /api/vocabularies/{vcode}/open
PUT /api/vocabularies/{vcode}/close
POST /api/vocabularies/{vcode}/concepts
GET|PUT|DELETE /api/vocabularies/{vcode}/concepts/{ccode}
PUT /api/vocabularies/{vcode}/concepts/{ccode}/commit
POST /api/vocabularies/{vcode}/concepts/{ccode}/merge
GET|PUT|DELETE /api/vocabularies/{vcode}/concepts/{ccode}
GET /api/concept-relations
GET /api/concept-search
PUT /api/concept-reindex

Technical integration

- **No direct connection** to a vocabulary server (like Skosmos)
- **Ingest** of vocabulary by uploading a **ttl-file** containing data in SKOS schema (see Swagger for details)
- Simple **curation** built-in…
- …but obstacles due to **missing end-points** and integration of simple SKOS
- Difference between **closed/open** vocabularies: open allows extensions by users but these needs approvement by curators – closed are freeze/read-only



GET /api/property-types Get all property types in pages

POST /api/property-types Create property type

POST /api/property-types/reorder Reorganize property type order

GET /api/vocabularies/{code} Get vocabulary for given code

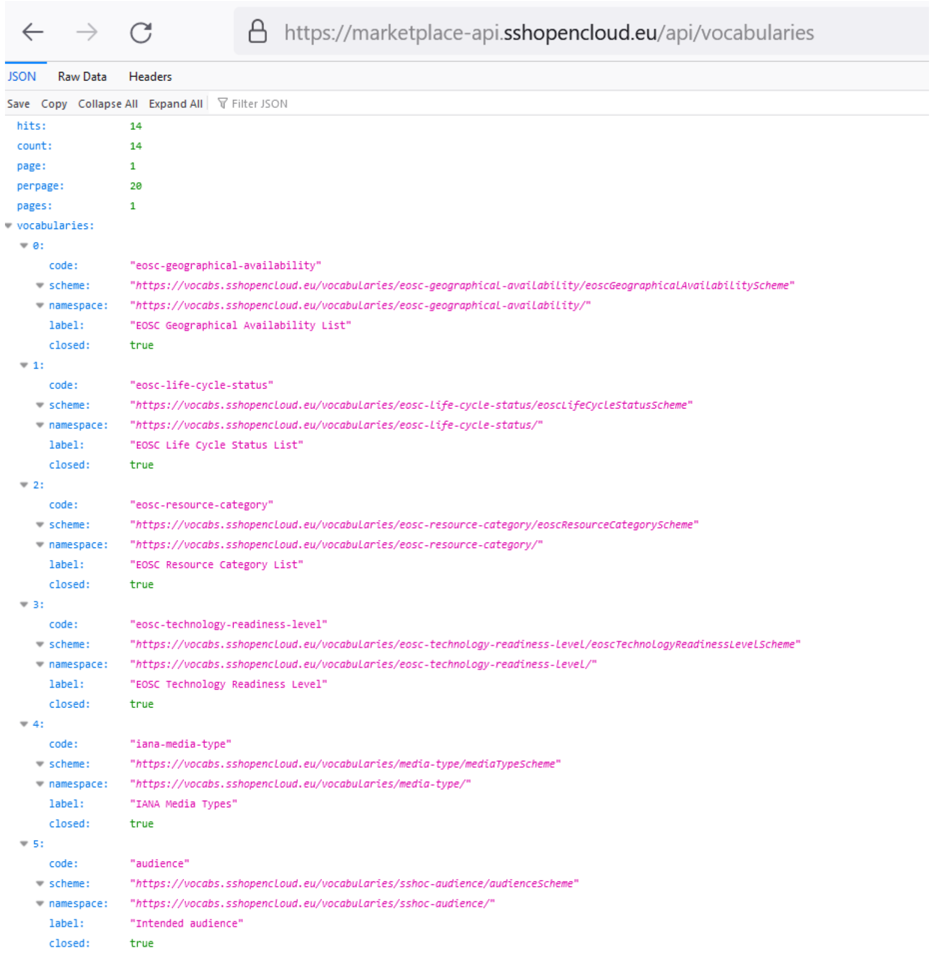
PUT /api/vocabularies/{code} Update vocabulary for given code and file

Parameters

Name	Description
code * required string *(path)*	code
ttl * required object *(query)*	

Development

- Currently 14 vocabularies: 13 **closed**, one **open** (= "sshoc-keyword")
- Connected to **properties**
- Declaration and definition of properties mostly derived from **ingestion sources**
- Some of the properties were identified as useful to be **based on a vocabulary holding concepts**
- Additionally, some of these concept properties qualified for **facets**

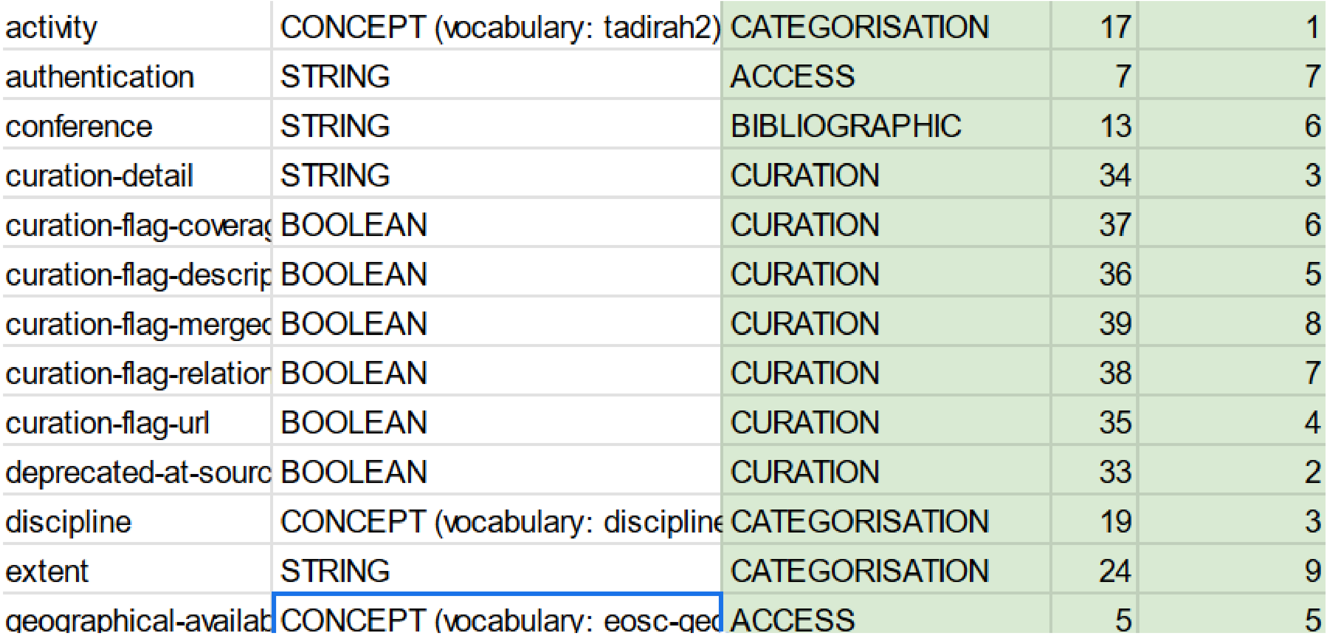


```
hits:           14
count:          14
page:           1
perpage:        20
pages:          1
▼ vocabularies:
  ▼ 0:
      code:         "eosc-geographical-availability"
    ▼ scheme:       "https://vocabs.sshopencloud.eu/vocabularies/eosc-geographical-availability/eoscGeographicalAvailabilityScheme"
    ▼ namespace:    "https://vocabs.sshopencloud.eu/vocabularies/eosc-geographical-availability/"
      label:        "EOSC Geographical Availability List"
      closed:       true
  ▼ 1:
      code:         "eosc-life-cycle-status"
    ▼ scheme:       "https://vocabs.sshopencloud.eu/vocabularies/eosc-life-cycle-status/eoscLifeCycleStatusScheme"
    ▼ namespace:    "https://vocabs.sshopencloud.eu/vocabularies/eosc-life-cycle-status/"
      label:        "EOSC Life Cycle Status List"
      closed:       true
  ▼ 2:
      code:         "eosc-resource-category"
    ▼ scheme:       "https://vocabs.sshopencloud.eu/vocabularies/eosc-resource-category/eoscResourceCategoryScheme"
    ▼ namespace:    "https://vocabs.sshopencloud.eu/vocabularies/eosc-resource-category/"
      label:        "EOSC Resource Category List"
      closed:       true
  ▼ 3:
      code:         "eosc-technology-readiness-level"
    ▼ scheme:       "https://vocabs.sshopencloud.eu/vocabularies/eosc-technology-readiness-level/eoscTechnologyReadinessLevelScheme"
    ▼ namespace:    "https://vocabs.sshopencloud.eu/vocabularies/eosc-technology-readiness-level/"
      label:        "EOSC Technology Readiness Level"
      closed:       true
  ▼ 4:
      code:         "iana-media-type"
    ▼ scheme:       "https://vocabs.sshopencloud.eu/vocabularies/media-type/mediaTypeScheme"
    ▼ namespace:    "https://vocabs.sshopencloud.eu/vocabularies/media-type/"
      label:        "IANA Media Types"
      closed:       true
  ▼ 5:
      code:         "audience"
    ▼ scheme:       "https://vocabs.sshopencloud.eu/vocabularies/sshoc-audience/audienceScheme"
    ▼ namespace:    "https://vocabs.sshopencloud.eu/vocabularies/sshoc-audience/"
      label:        "Intended audience"
      closed:       true
```

Development

Choosing a suiting vocabulary based on pragmatically-driven workflow:

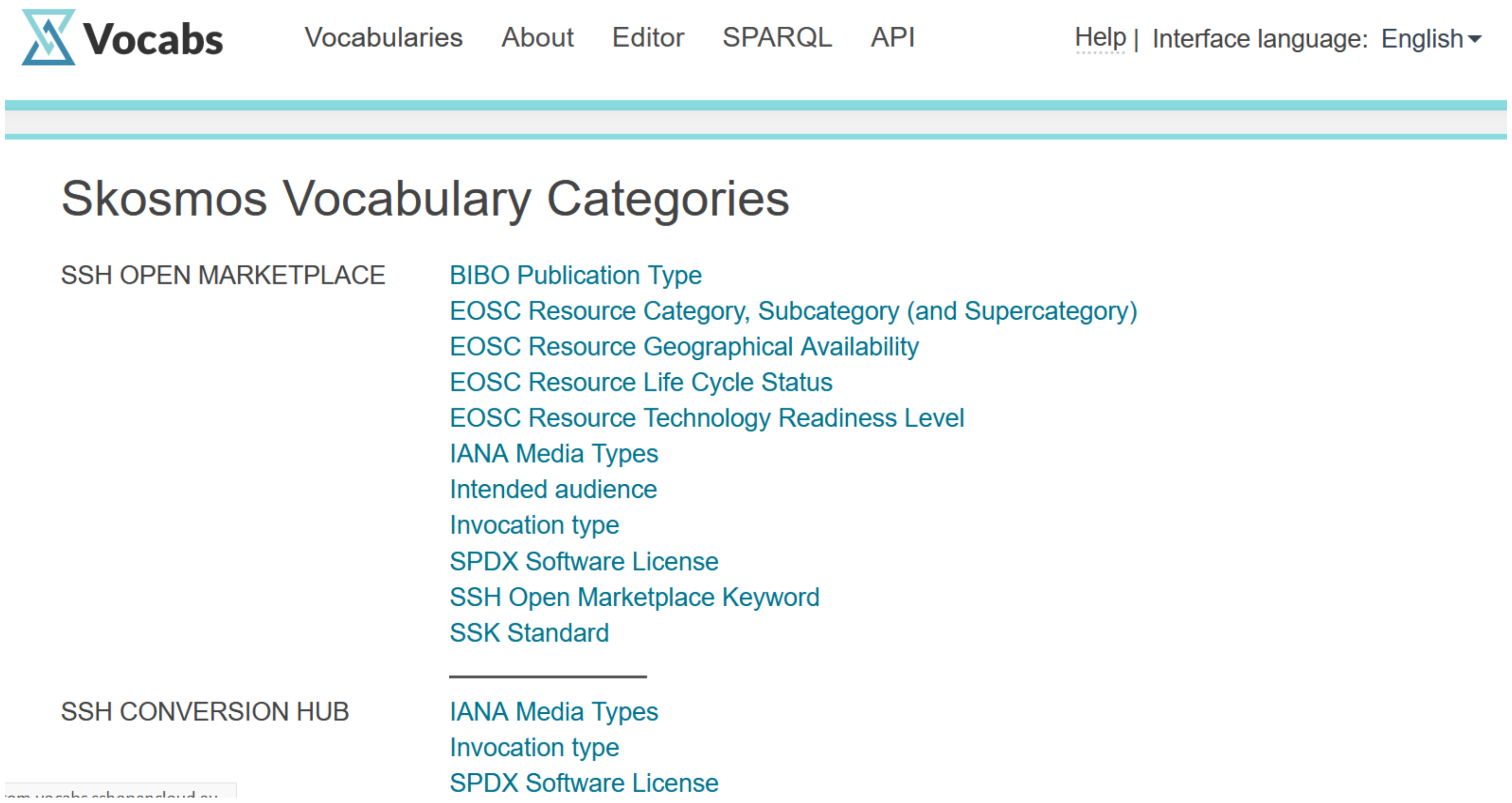
1. Existence of an EOSC vocabulary/ resource profile
2. Availability of a vocabulary at the source of ingestion
3. Finding a proper vocabulary at DARIAH vocabulary server
4. Looking for a proper vocabulary somewhere else, e.g., via BARTOC
5. Creating a dedicated vocabulary



activity	CONCEPT (vocabulary: tadirah2)	CATEGORISATION	17	1
authentication	STRING	ACCESS	7	7
conference	STRING	BIBLIOGRAPHIC	13	6
curation-detail	STRING	CURATION	34	3
curation-flag-coverag	BOOLEAN	CURATION	37	6
curation-flag-descrip	BOOLEAN	CURATION	36	5
curation-flag-mergec	BOOLEAN	CURATION	39	8
curation-flag-relatior	BOOLEAN	CURATION	38	7
curation-flag-url	BOOLEAN	CURATION	35	4
deprecated-at-sourc	BOOLEAN	CURATION	33	2
discipline	CONCEPT (vocabulary: discipline	CATEGORISATION	19	3
extent	STRING	CATEGORISATION	24	9
geographical-availab	CONCEPT (vocabulary: eosc-geo	ACCESS	5	5

Development

- SSH vocabs commons as **additional place** to communicate the used vocabularies in SSHOMP
- Contains all vocabularies that are not on DARIAH vocabs
- **Overlap of outcomes** within SSHOC project, e.g., IANA Media Types used for SSHOMP and Conversion Hub, "invocation type" developed for Conversion Hub and used for SSHOMP

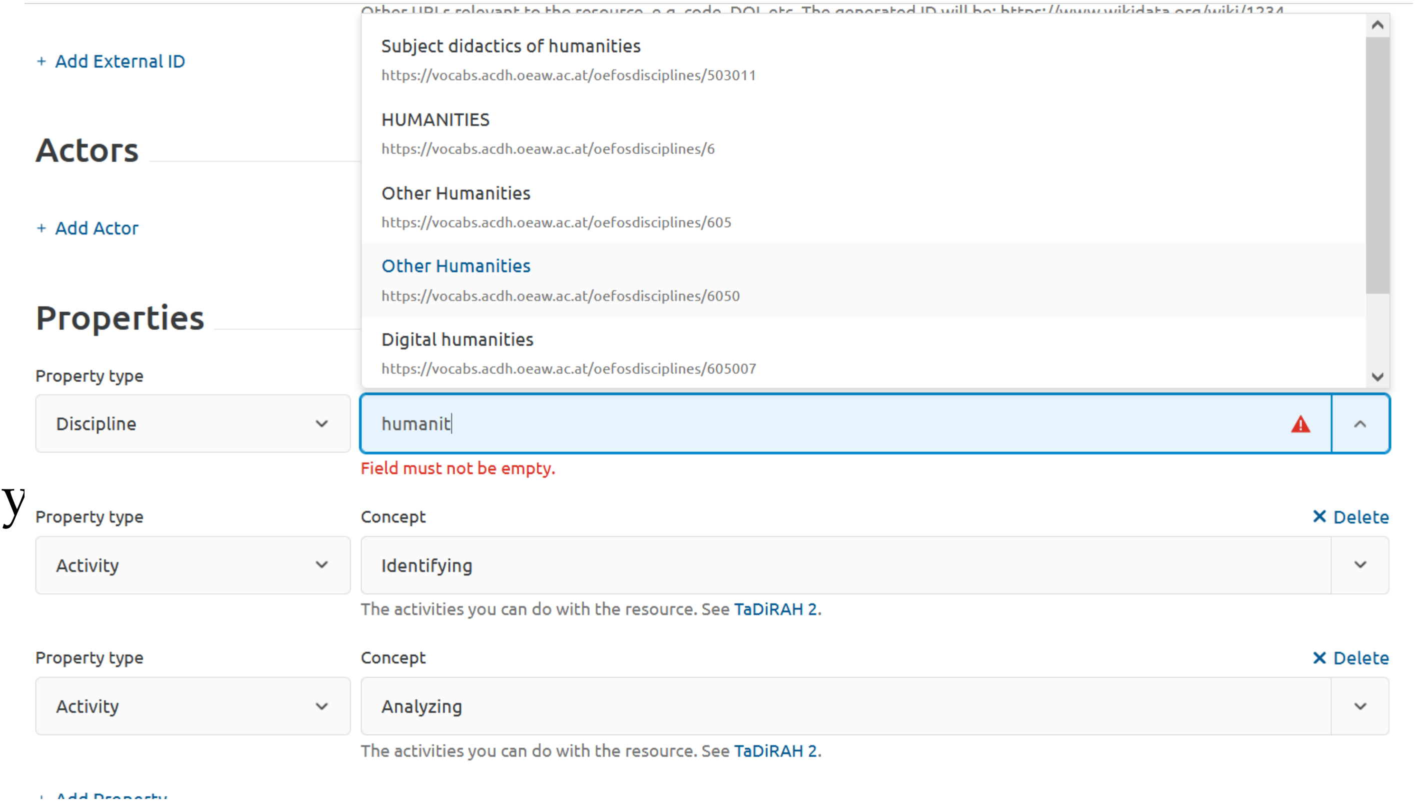


Skosmos Vocabulary Categories

SSH OPEN MARKETPLACE	BIBO Publication Type EOSC Resource Category, Subcategory (and Supercategory) EOSC Resource Geographical Availability EOSC Resource Life Cycle Status EOSC Resource Technology Readiness Level IANA Media Types Intended audience Invocation type SPDX Software License SSH Open Marketplace Keyword SSK Standard
SSH CONVERSION HUB	IANA Media Types Invocation type SPDX Software License

UX for data input

- User experience is challenging, especially for **input forms**
- **Many** properties, many concepts
- **Autocomplete** helpful but not fully satisfying
- **Hierarchy** not shown in frontend
- **Navigation** needs more clever approach, not easy to implement
- Current approach focus **dynamic development of properties**: probably at cost of usability for using complex vocabularies



Other URLs relevant to the resource, e.g. code, DOI, etc. The generated ID will be: https://www.wikidata.org/wiki/1234

+ Add External ID

Actors

+ Add Actor

Properties

Property type: Discipline

humanit

Subject didactics of humanities
https://vocabs.acdh.oeaw.ac.at/oefosdisciplines/503011

HUMANITIES
https://vocabs.acdh.oeaw.ac.at/oefosdisciplines/6

Other Humanities
https://vocabs.acdh.oeaw.ac.at/oefosdisciplines/605

Other Humanities
https://vocabs.acdh.oeaw.ac.at/oefosdisciplines/6050

Digital humanities
https://vocabs.acdh.oeaw.ac.at/oefosdisciplines/605007

Field must not be empty.

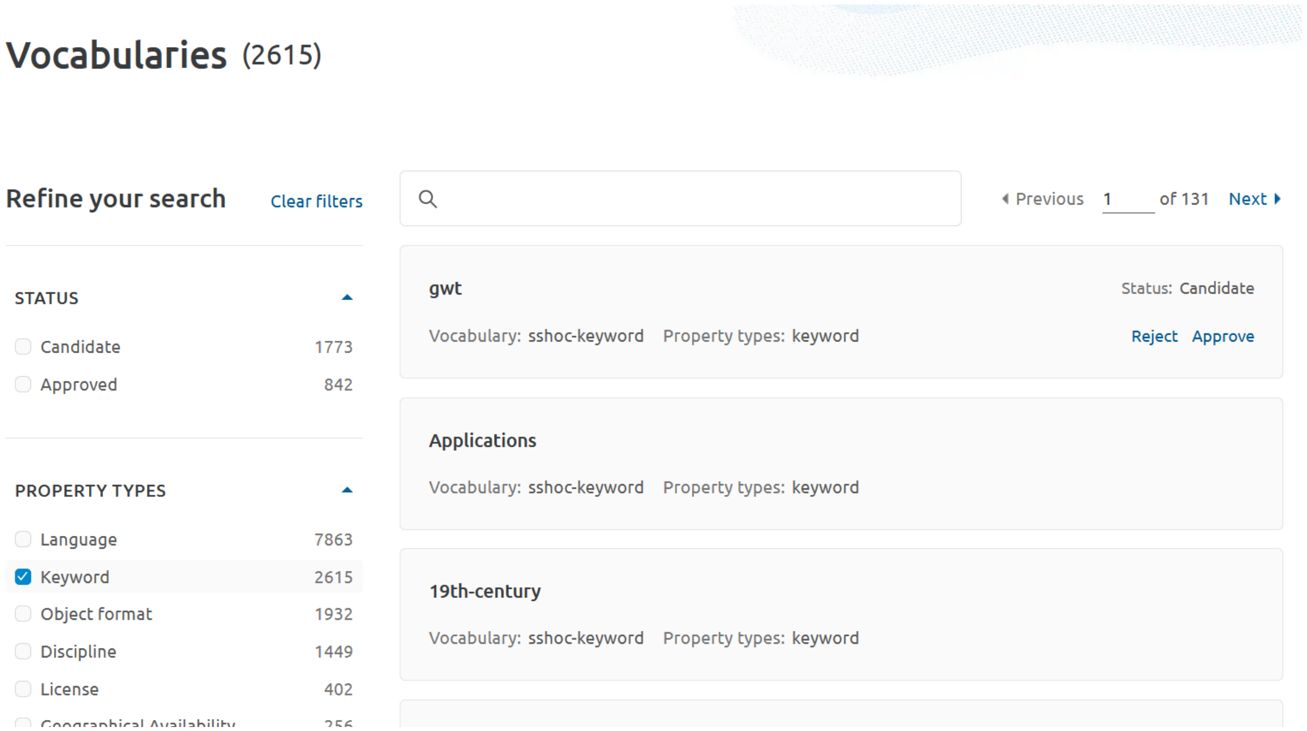
Property type: Activity
Concept: Identifying
✕ Delete
The activities you can do with the resource. See TaDiRAH 2.

Property type: Activity
Concept: Analyzing
✕ Delete
The activities you can do with the resource. See TaDiRAH 2.

+ Add Property

Curation

- **Creation** of vocabularies outsourced to **external tools**, e.g., ACDH-CH vocabs editor
- Closed vocabularies curation needs **ttl dump** and full upload
- **Open vocabularies** curation partly implemented: **candidates** are concepts proposed by users and to be approved by moderators
- **Mapping logic in ingestion pipelines**: created many duplicates esp. for vocabulary sshoc-keyword > identify and merge **duplicates**



Vocabularies (2615)

Refine your search Clear filters

◂ Previous 1 of 131 Next ▸

STATUS

□ Candidate	1773
□ Approved	842

PROPERTY TYPES

□ Language	7863
☑ Keyword	2615
□ Object format	1932
□ Discipline	1449
□ License	402
□ Geographical Availability	256

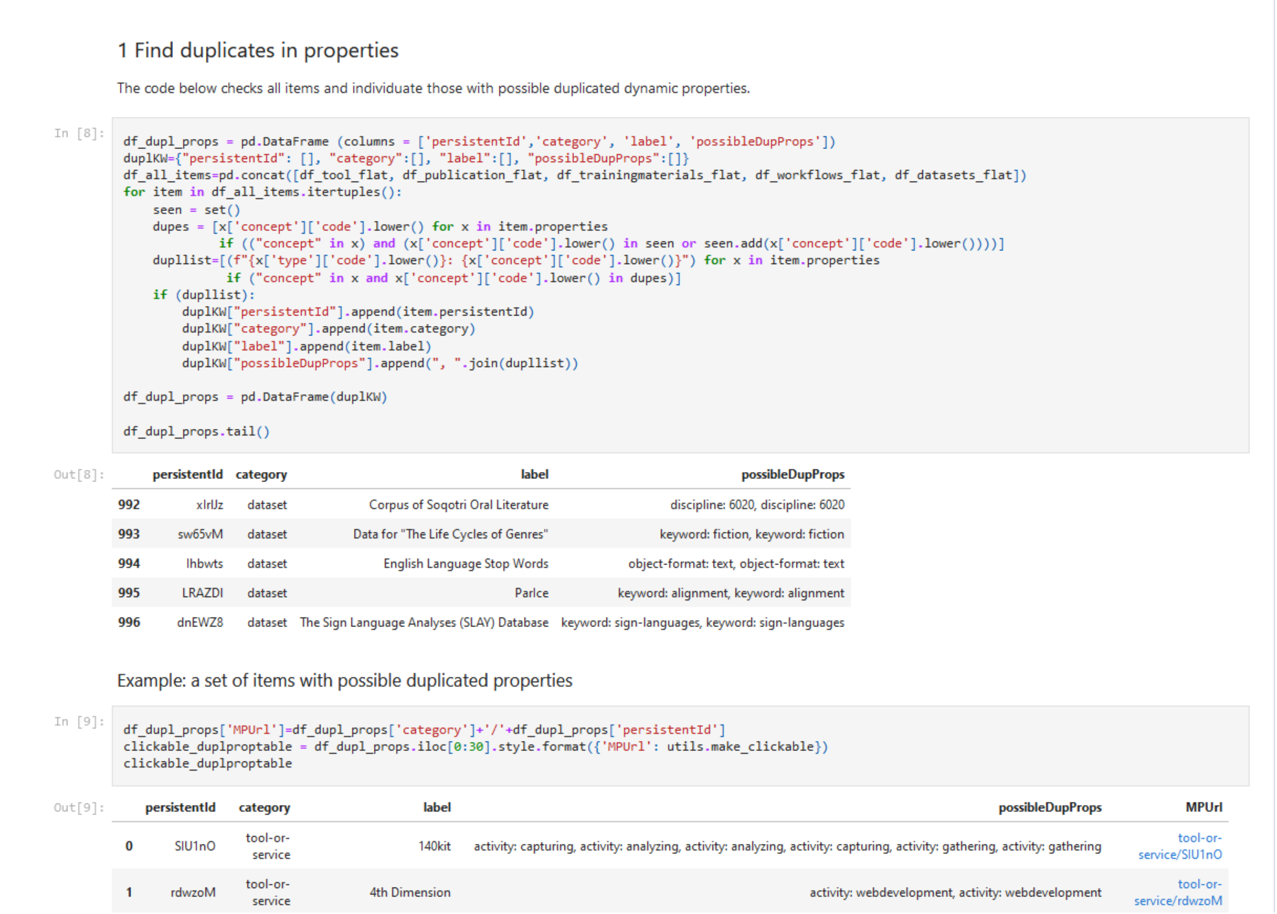
gwt
Status: Candidate
Vocabulary: sshoc-keyword Property types: keyword
Reject Approve

Applications
Vocabulary: sshoc-keyword Property types: keyword

19th-century
Vocabulary: sshoc-keyword Property types: keyword

Curation

- **Complex** (= many steps to proceed) **curation work** not possible alone with frontend
- API allows to dock into SSHOMP with **other tools**: Python notebooks are used for extended curation
- Curation **team**: Cesare Concordia, Laure Barbot, Martin Kirnbauer
- Collection of **notebooks**: https://github.com/SSHOC/marketplace-curation



1 Find duplicates in properties

The code below checks all items and individuate those with possible duplicated dynamic properties.

In [8]:

```
df_dupl_props = pd.DataFrame (columns = ['persistentId','category', 'label', 'possibleDupProps'])
duplKW={"persistentId": [], "category":[], "label":[], "possibleDupProps":[]}
df_all_items=pd.concat([df_tool_flat, df_publication_flat, df_trainingmaterials_flat, df_workflows_flat, df_datasets_flat])
for item in df_all_items.itertuples():
    seen = set()
    dupes = [x['concept']['code'].lower() for x in item.properties
             if (("concept" in x) and (x['concept']['code'].lower() in seen or seen.add(x['concept']['code'].lower())))]
    dupllist=[(f"{x['type']['code'].lower()}: {x['concept']['code'].lower()}") for x in item.properties
              if ("concept" in x and x['concept']['code'].lower() in dupes)]
    if (dupllist):
        duplKW["persistentId"].append(item.persistentId)
        duplKW["category"].append(item.category)
        duplKW["label"].append(item.label)
        duplKW["possibleDupProps"].append(", ".join(dupllist))

df_dupl_props = pd.DataFrame(duplKW)

df_dupl_props.tail()
```

Out[8]:

	persistentId	category	label	possibleDupProps
992	xlrIJz	dataset	Corpus of Soqotri Oral Literature	discipline: 6020, discipline: 6020
993	sw65vM	dataset	Data for "The Life Cycles of Genres"	keyword: fiction, keyword: fiction
994	lhbwts	dataset	English Language Stop Words	object-format: text, object-format: text
995	LRAZDI	dataset	Parlce	keyword: alignment, keyword: alignment
996	dnEWZ8	dataset	The Sign Language Analyses (SLAY) Database	keyword: sign-languages, keyword: sign-languages

Example: a set of items with possible duplicated properties

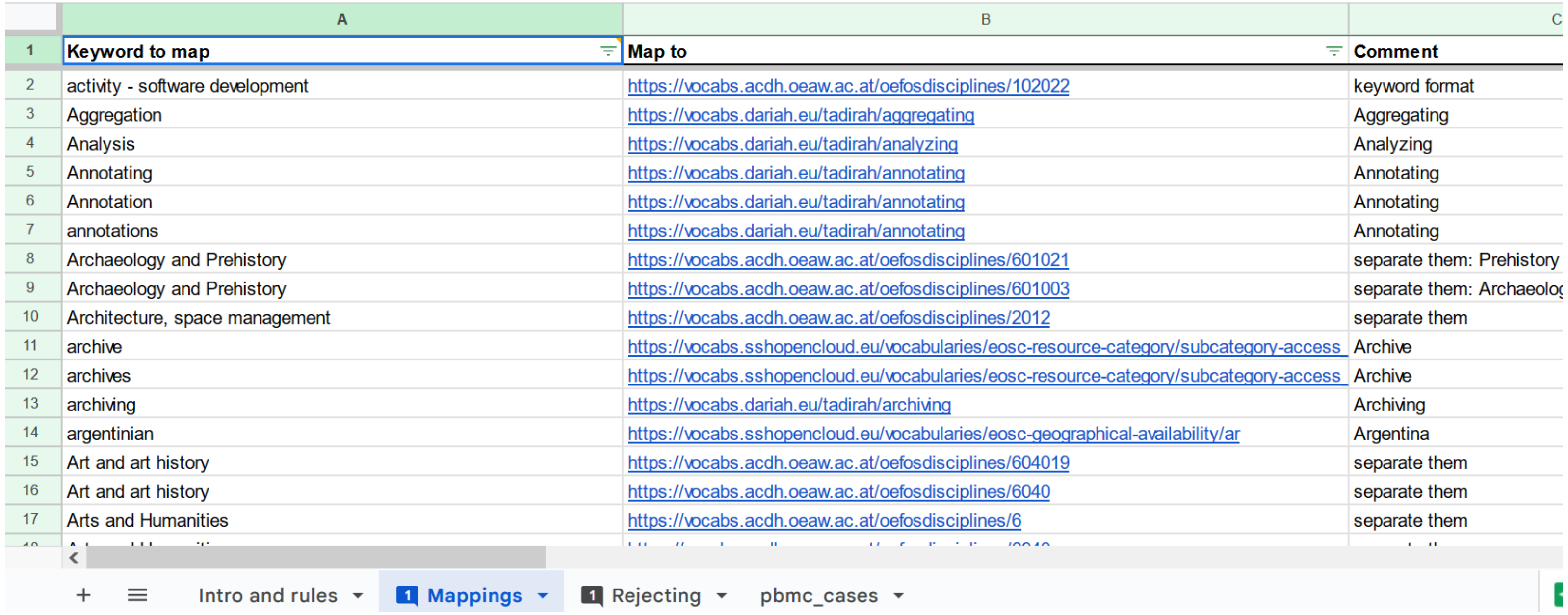
In [9]:

```
df_dupl_props['MPUrl']=df_dupl_props['category']+'/'+df_dupl_props['persistentId']
clickable_duplproptable = df_dupl_props.iloc[0:30].style.format({'MPUrl': utils.make_clickable})
clickable_duplproptable
```

Out[9]:

	persistentId	category	label	possibleDupProps	MPUrl
0	SIU1nO	tool-or-service	140kit	activity: capturing, activity: analyzing, activity: analyzing, activity: capturing, activity: gathering, activity: gathering	tool-or-service/SIU1nO
1	rdwzoM	tool-or-service	4th Dimension	activity: webdevelopment, activity: webdevelopment	tool-or-service/rdwzoM

Curation



	Keyword to map	Map to	Comment
2	activity - software development	https://vocabs.acdh.oeaw.ac.at/oefosdisciplines/102022	keyword format
3	Aggregation	https://vocabs.dariah.eu/tadirah/aggregating	Aggregating
4	Analysis	https://vocabs.dariah.eu/tadirah/analyzing	Analyzing
5	Annotating	https://vocabs.dariah.eu/tadirah/annotating	Annotating
6	Annotation	https://vocabs.dariah.eu/tadirah/annotating	Annotating
7	annotations	https://vocabs.dariah.eu/tadirah/annotating	Annotating
8	Archaeology and Prehistory	https://vocabs.acdh.oeaw.ac.at/oefosdisciplines/601021	separate them: Prehistory
9	Archaeology and Prehistory	https://vocabs.acdh.oeaw.ac.at/oefosdisciplines/601003	separate them: Archaeolog
10	Architecture, space management	https://vocabs.acdh.oeaw.ac.at/oefosdisciplines/2012	separate them
11	archive	https://vocabs.sshopencloud.eu/vocabularies/eosc-resource-category/subcategory-access	Archive
12	archives	https://vocabs.sshopencloud.eu/vocabularies/eosc-resource-category/subcategory-access	Archive
13	archiving	https://vocabs.dariah.eu/tadirah/archiving	Archiving
14	argentinian	https://vocabs.sshopencloud.eu/vocabularies/eosc-geographical-availability/ar	Argentina
15	Art and art history	https://vocabs.acdh.oeaw.ac.at/oefosdisciplines/604019	separate them
16	Art and art history	https://vocabs.acdh.oeaw.ac.at/oefosdisciplines/6040	separate them
17	Arts and Humanities	https://vocabs.acdh.oeaw.ac.at/oefosdisciplines/6	separate them

- Includes a lot of different work regarding **data quality**
- Idea of **monitoring** if there is the need for a new closed vocabulary by looking at evolution of open vocabulary sshoc-keyword: **identify** possible **new concept properties/vocabularies**
- **Closed** vocabularies good to handle, but not always easy to **map**
- **Open** vocabulary allows **flexibility** and **user interaction**, but needs strong and as good as it gets automated curation
- **Notebooks** for doing **automated curation** – SSH vocabs commons as exploratory tool to support mapping/choosing concepts (not implemented in SSHOMP backend)

Analysis

- Analysis of **usage** (via SQL), state of 24/03/2023: 14 **vocabularies** connected to 14 **properties** having 15.012 **concepts**, and 6.574 **items** (including steps) having 37.614 **applied properties** where 25.000 applied properties having values that are **concepts of a vocabulary** (= 66,5 %)
- The **most used concepts for properties** are: 2.731 "Conference" (publication-type), 1.078 "eng" (iso-639-3), 620 "analyzing" (tadirah2), 541 "CC-BY-4.0" (software license), 489 "webApplication" (invocationType) and 347 "Other" (sshoc-keyword)
- **Most used concepts from open vocabulary** "sshoc-keyword": 347 "Other", 193 "american", 180 "tokenised", 169 "search", 162 "Spoken+corpora", 162 "natural-language-processing"
- Looking at **user generated vs. ingested ones**: 286 items **created by users** having 2.627 properties (= 9,2 per item) where 2.201 are concept properties (= 83,8 % = 7,7 per item) with 678 being **keyword concepts** (= 30,8%), vs. 6.288 items **ingested** having 34.987 properties (= 5,6 per item) where 22.799 are concept properties (= 65,2 % = 3,6 per item) being 8.590 **keyword concepts** (= 37,7 %)

Experiences

(Some) **advantages** of vocabularies in SSHOMP

- Higher data quality & more structured data
- Potential to find overlaps with other data collections
- Attempt to collect (meta) vocabularies for the SSH domain
- Curation of vocabularies possible
- Aiming for FAIRness of vocabulary management
- Discovery works well with facets

and (some) **Disadvantages**

- Not easy to fill out input forms
- Mapping costs a lot of time and needs many concessions
- Only a tailored view on SSH domain based on data model of SSHOMP
- Curation coding centric due to notebooks
- FAIRness of vocabulary management only on a very basic level: findability is a problem, and re-useability does not happen often
- More ways needed to explore next to facets

Problems

- Often no **definitions** or under-specified information about concepts
- Tendency to prefer **simple vocabularies** in UX, e.g., flat hierarchy
- Mission statement of SSHOMP and pragmatic approach to find proper vocabularies does only **reflect a specific part** of SSH domain
- Technical setup is not focused on a **full vocabulary management workflow**, e.g., Skosmos is separated from SSHOMP API
- SSHOMP API needs to be **extended for a better vocabulary experience**, e.g., search for persistent identifiers of concepts does not work well
- Use of Python notebooks because SSHOMP frontend did not adopt a full **vocabulary management system**
- Creation/updates of vocabulary needs to be handled externally: subsequent updates not easy

Challenges

- Establish a **metadata scheme for describing the used vocabularies**: inject information about aim of SSHOMP, reasons for deciding to use this vocabulary, adaptions of vocabulary, etc. => possible in full text field but we like to have it in a structured way
- Clever **vocabulary handling for mappings** of new sources and creation of properties
- Establish different **workflows to handle** vocabulary management that deals with: **forking** vocabularies, using only **subset** of vocabularies, **merging** vocabularies, **connecting** concepts
- Keep track with **updates of vocabularies** that are re-used at source
- **Connect vocabulary concepts** from SSHOMP domain with other domains – use SKOS power (closeMatch, exactMatch, broadMatch, …) on concepts or on an in-between-vocabulary
- Curation in general due to the technical constraints: extension and documention of curation
- keywords tend to be **added again and again** => map from open vocabulary to closed ones

Take aways

- Good **vocabulary management** helps to create data with **good quality**/structured information
- **Well-designed combination** of frontend, backend, vocabulary creation system, vocabulary representation system, mapping tools, ingestion pipelines, data input is necessary
- **Dissemination of vocabularies** important: how to address the community of the domain?
- Update information of external re-used vocabularies not always easy to get: how to establish such **communication channels**?
- Item history implemented in SSHOMP but not vocabulary/concept history: would be also good information => needs **extra metadata schemas** next to SKOS, e.g., PROV-O
- **Finding vocabularies** and deciding which ones **to use** is challenging > interestingly many (tacit) vocabulary information is not published/available in SKOS , e.g., simple combo boxes
- **SSH vocabs commons** as a place to solve some of this issues? **Mission statement** necessary!

Thank you for your attention!
Time for discussion …

klaus.illmayer@oeaw.ac.at