



universität
wien



COAR Resource Type Vocabulary in the Classification Server of Phaidra

Sandor Kopacsi

Computer Center, University of Vienna

COAR Annual Meeting 2017
May 8-10, 2017, Venice, Italy

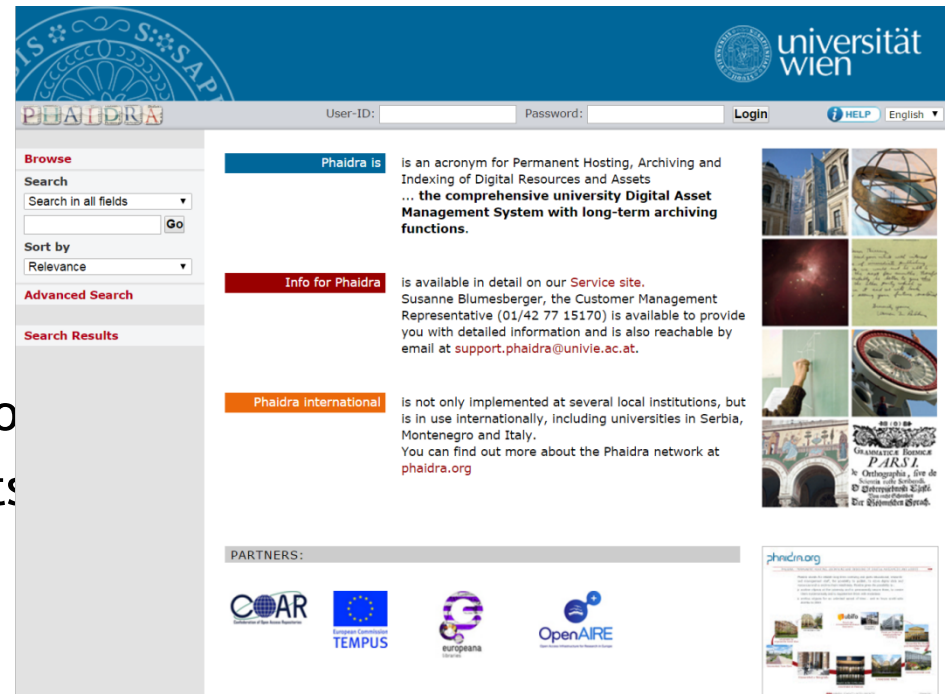
PHAIDRA = Permanent Hosting, Archiving and Indexing of Digital Resources and Assets

<https://phaidra.univie.ac.at/>

- reliable long-term preservation system of University of Vienna
- gives the possibility to the educational, research- and management staff
 - to publish, store, archive, access and retrieve
 - digital data and resources
 - taking copy rights into account

Fields of use

- research (e.g. articles, videos, audio)
- learning & teaching (learning objects)
- administration (e.g. forms)

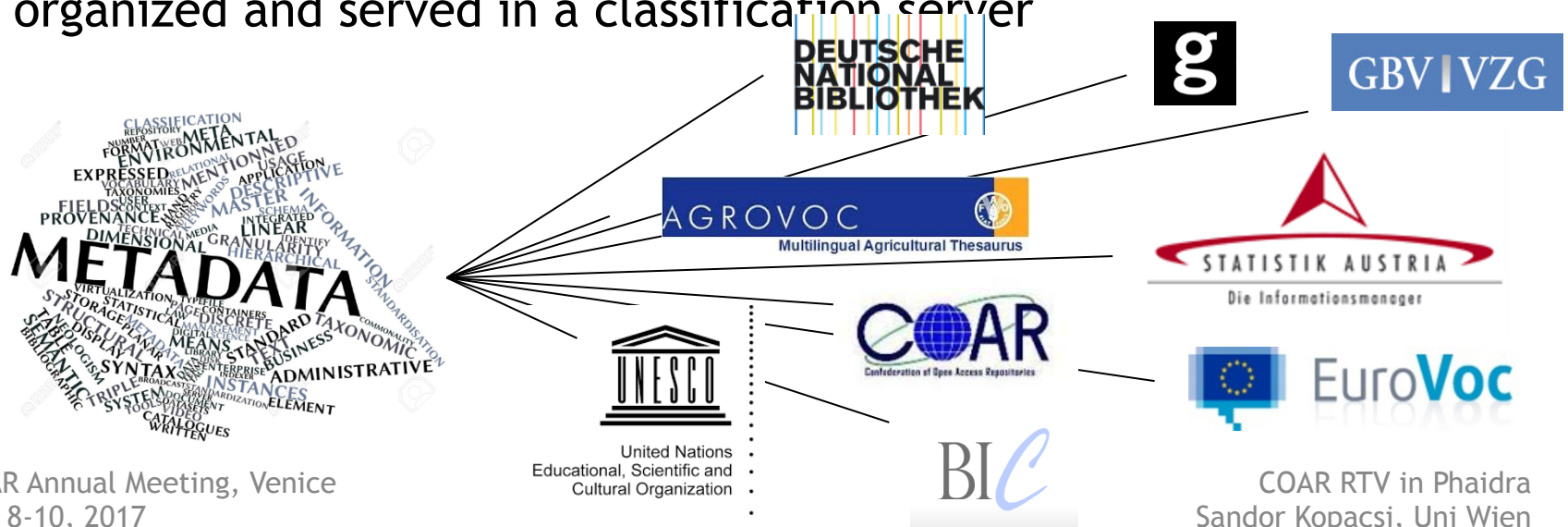


The screenshot shows the Phaidra website interface. At the top, there is a blue header with the University of Vienna logo and the text 'universität wien'. Below the header, there is a search bar with 'User-ID:' and 'Password:' fields, and a 'Login' button. The main content area is divided into several sections:

- Browse:** A search bar with 'Search in all fields' and a 'Go' button. Below it, there are options for 'Sort by' and 'Relevance'.
- Advanced Search:** A section for more detailed search options.
- Search Results:** A section for displaying search results.
- Phaidra is:** A section describing Phaidra as an acronym for Permanent Hosting, Archiving and Indexing of Digital Resources and Assets, and as the comprehensive university Digital Asset Management System with long-term archiving functions.
- Info for Phaidra:** A section providing contact information for Susanne Blumesberger, the Customer Management Representative, and a link to the service site.
- Phaidra International:** A section mentioning that Phaidra is implemented at several local institutions and is also used internationally, including universities in Serbia, Montenegro, and Italy.
- PARTNERS:** A section listing partner organizations, including COAR, European Commission TEMPUS, europeana, and OpenAIRE.

Necessity of a Classification Server

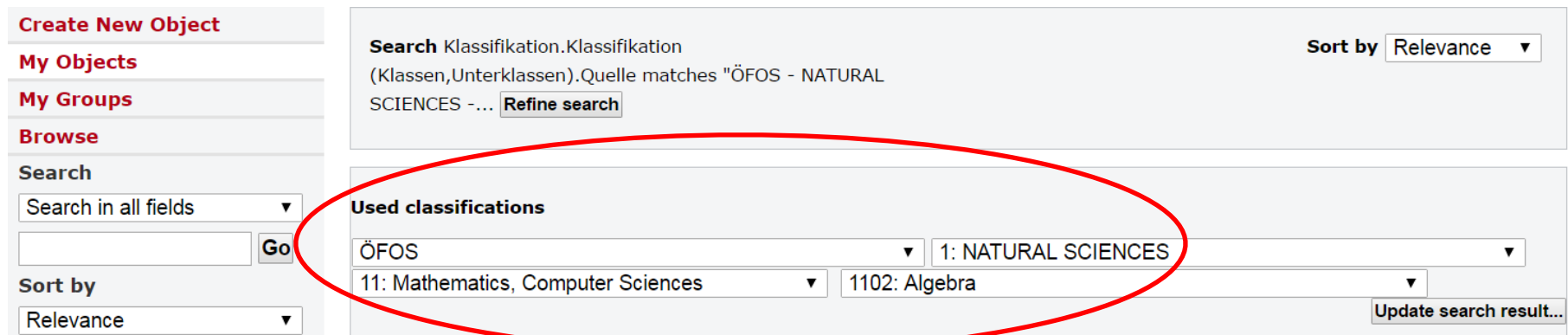
- the objects can be more easily found, if we use well-defined metadata by storing and searching
- selecting metadata terms from pre-defined, controlled vocabularies and classifications
- classifications in many topics are already available (e.g. Getty, AGROVOC, Eurovoc, ÖFOS, COAR RTV).
- our relevant vocabularies and classifications can be stored, organized and served in a classification server



Goals of the Classification Server

- metadata organisation
- metadata access
 - when the user **ingests new items** to the preservation system, and wants to assign metadata to it,
 - when the user **browses or searches** in the preservation system for items supplied with metadata terms.

The search/browse/ingest is easier from controlled vocabularies, or other classifications in a (hierarchically) arranged structure.



The screenshot shows the Classification Server interface. On the left, there are navigation links: "Create New Object", "My Objects", "My Groups", and "Browse". Below these is a search section with a dropdown menu set to "Search in all fields", a search input field, and a "Go" button. Below the search section is a "Sort by" dropdown menu set to "Relevance". The main content area shows search results for "Klassifikation.Klassifikation" with a "Sort by" dropdown set to "Relevance". Below the search results is a section titled "Used classifications" which is circled in red. This section contains two dropdown menus: the first is set to "ÖFOS" and the second is set to "1: NATURAL SCIENCES". Below these are two more dropdown menus: the first is set to "11: Mathematics, Computer Sciences" and the second is set to "1102: Algebra". At the bottom right of the "Used classifications" section is an "Update search result..." button.

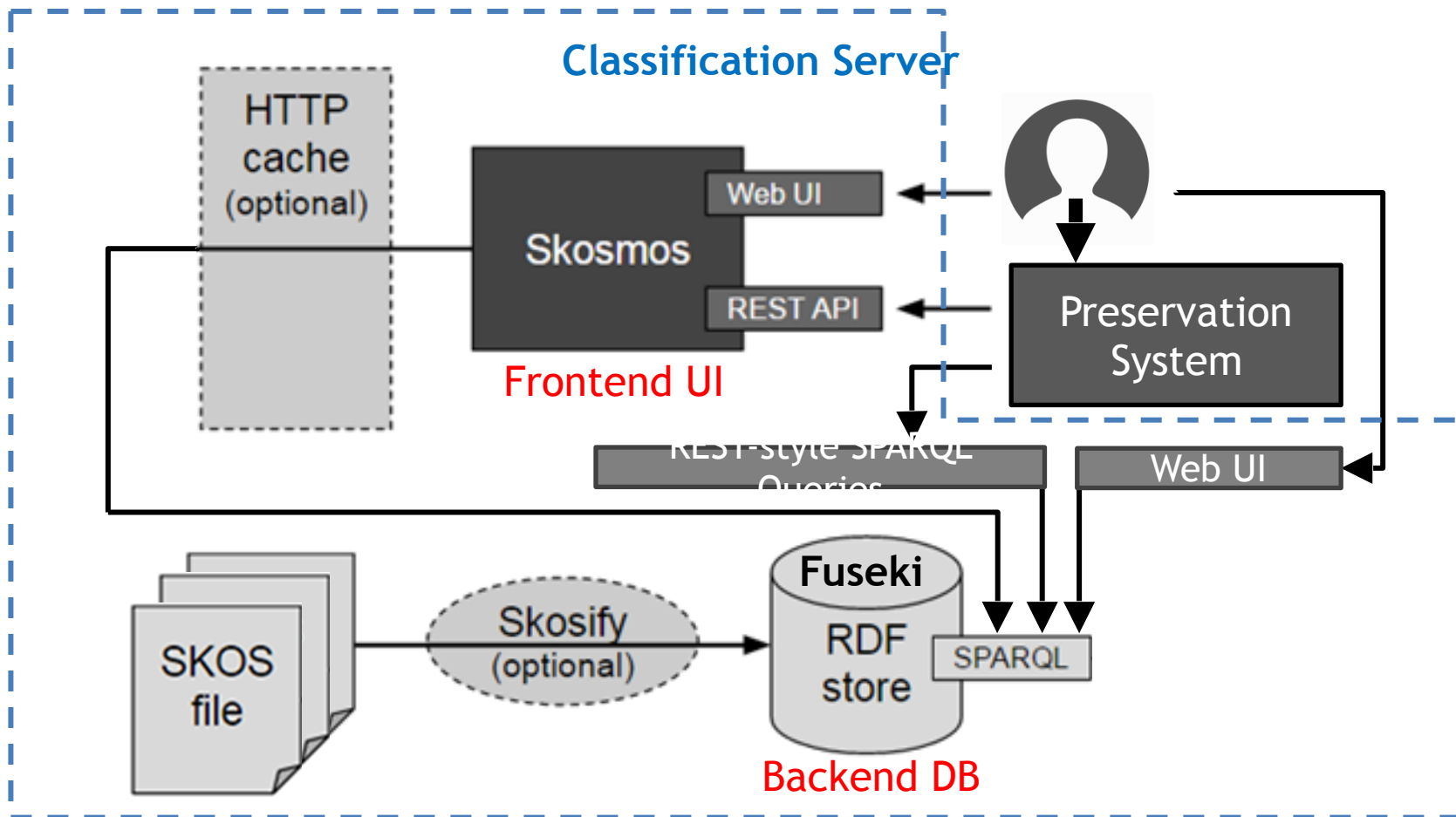


Implementation in Skosmos with Jena Fuseki triple store



- **Open source** web application for browsing classifications.
- Developed by the **National Library of Finland**.
- Built on the basis of prior development: **ONKI Light**.
- Classifications are stored in RDF triple stores accessed via **SPARQL endpoints**.
- Recommended Backend DB: **Jena Fuseki triple store**.
- Provides a **multilingual user interface**.
- Delivers **REST API** to access through HTTP.

System Architecture

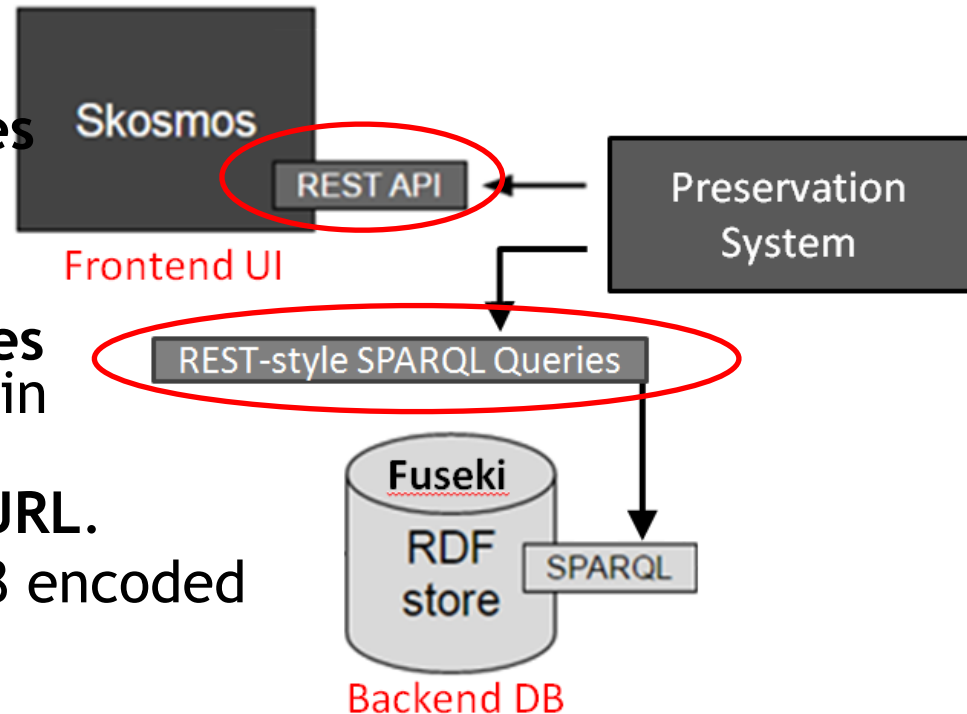


Connection the Preservation System to the Classification Server

The connection will be realised using the

- REST API of Skosmos and/or
- REST-style SPARQL Queries of Jena Fuseki

- These are **read-only interfaces** over HTTP to the data stored in the Classification Server.
- Requests can be built in the **URL**.
- The **returned data** is in UTF-8 encoded **JSON-LD** format.



Skosmos

Vocabularies About Feedback Help | auf Deutsch in Italiano na srpskom

from all English | x Search



Welcome to the
Classification Server of
PHAIDRA

PHAIDRA is the
comprehensive Digital
Asset Management System
of the University of Vienna
with long-term archiving
functions.



internal
vocabularies

Classification Categories

GENERAL ONLINE CLASSIFICATIONS AGROVOC - Multilingual agricultural thesaurus
Eurovoc - EU's multilingual thesaurus
STW Thesaurus for Economics
UNESCO Thesaurus

GENERAL LOCAL CLASSIFICATIONS COAR Resource Type Vocabulary v1.1
GND Subject Categories
ÖFOS 2012 (Juli 2015)
STW Thesaurus for Economics
UNESCO Thesaurus

GENERAL LOCAL CLASSIFICATIONS FROM PHAIDRA
ACM 1998 - The ACM Computing Classification System [1998 Version]
Basisklassifikation
BIC Standard Subject Categories
BIC Standard Subject Qualifiers
Dewey Decimal Classification
ÖFOS 2002
Physics and Astronomy Classification Scheme

PHAIDRA'S LOCAL CLASSIFICATIONS Phaidra's controlled vocabularies
Phaidra's custom classifications

search in all
vocabularies

external
vocabularies

COAR Resource Type Vocabulary v1.1

Content language English ▾ × Search

Alphabetical Hierarchy New

- cartographic material
- dataset
- image
- interactive resource
- other
- software
- sound
- text
- workflow

Vocabulary information

TITLE COAR Resource Type Vocabulary v1.1

TYPE <http://www.w3.org/2004/02/skos/core#ConceptScheme>

URI http://purl.org/coar/resource_type

Resource counts by type

Type	Count
Concept	58

Term counts by language

Language	Preferred terms	Alternate terms	Hidden terms
Catalan	56	6	0
German	56	17	0
English	58	31	0
Spanish	57	133	0
French	56	49	0
Italian	56	26	0
Japanese	56	0	0
Dutch	54	13	0
Portuguese	56	83	0
Russian	56	293	0
Turkish	51	6	0
Chinese	58	144	0

Skosmos

Vocabularies About Feedback Help | auf Deutsch in Italiano na srpskom

COAR Resource Type Vocabulary v1.1

Content language English Search

Alphabetical Hierarchy **New**

- text
 - annotation
 - bibliography
 - book
 - conference object
 - lecture
 - letter
 - musical notation
 - patent
 - periodical
 - journal
 - contribution to journal
 - data paper
 - editorial
 - journal article**
 - research article
 - review article
 - letter to the editor
 - preprint
 - report
 - research proposal
 - review
 - technical documentation
 - thesis
 - working paper

text > periodical > journal > contribution to journal > journal article

PREFERRED TERM

journal article

DEFINITION

i An article on a particular topic and published in a journal issue. (adapted from fabio)

BROADER CONCEPT

[contribution to journal](#)

NARROWER CONCEPTS

[research article](#)
[review article](#)

RELATED CONCEPTS

<http://purl.org/eprint/type/JournalItem>
<http://purl.org/eprint/type/SubmittedJournalArticle>

ALTERNATIVE LABEL

article
journal articles
paper

IN OTHER LANGUAGES

article de revista	Catalan
<i>article</i>	
学术论文	Chinese
期刊論文	
期刊文章	
學術論文	
論文	
期刊论文	
论文	
文章	
wetenschappelijk artikel	Dutch
article	French
<i>article scientifique</i>	
<i>article de recherche</i>	
Wissenschaftlicher Artikel	German
articolo in rivista	Italian
<i>articolo</i>	

Problems with COAR RTV

- labels using SKOS-XL properties, but Skosmos doesn't support SKOS-XL currently
- imported SKOS-XL version doesn't contain the `skos:narrower` tag, just the `skos:broader`, which is not enough for Skosmos.
- `dct:created` data is in `xsd:dateTime` format (e.g. "2015-06-03T20:06:48Z", that format is not handled by Skosmos

COAR RTV: conversion SKOS-XL to SKOS core

using owlart converter

1. download from <https://bitbucket.org/art-uniroma2/owlart/download>
2. copy files to be converted
3. create an output file
4. create config
 - set baseuri
 - namespaces
 - defaultSchem
5. Run

SPARQL Update queries

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
PREFIX skosxl: <http://www.w3.org/2008/05/skos-xl#>
INSERT {
  ?c skos:prefLabel ?pref .
  ?c skos:altLabel ?alt .
  ?c skos:hiddenLabel ?hidden .
  ?c skos:definition ?def .
}
WHERE {
  { ?c skosxl:prefLabel/skosxl:literalForm ?pref }
  UNION
  { ?c skosxl:altLabel/skosxl:literalForm ?alt }
  UNION
  { ?c skosxl:hiddenLabel/skosxl:literalForm ?hidden }
  UNION
  { ?c skos:definition/rdf:value ?def }
}
```

COAR RTV: missing skos:narrower tag

- we have used the Skosify tool (<http://demo.seco.tkk.fi/skosify/skosify>), but it creates `skos:narrower` as a pair of `skos:broadMatch`, wich is not correct
- Trick: renaming `skos:broadMatch` before skosifying and naming back afterwards

COAR RTV: xsd:dateTime to xsd:date format

- `dct:created` info is displayed in the „New” (earlier „Changes”) tab of Skosmos
- but `xsd:dateTime` is not processed („New” tab remains empty)
- Conversion from
`"2015-06-03T20:06:48Z"^^xsd:dateTime`
to `"2015-06-03"^^xsd:date`
- losses of information, but displayed

Conclusions and future plans

- we have successfully completed our research objectives:
 - to collect some available methods and tools for classification,
 - to implement a Classification Server
- Skosmos and Jena Fuseki seemed to be a good choice, despite of the difficulties during the implementation, as well as with the upload of certain classifications.
- there are hundreds of visits since the official launch of the Classification Server
- the general online external classifications (AGROVOC, Eurovoc, STW and UNESCO) are using currently the SPARQL server of the National Library of Finland
- in the near future we are going to redirect these external requests to their original data source
- the connection between the Classification Server and Phaidra is under development at the moment