State of BARTOC.org

Basic Register of Thesauri, Ontologies & Classifications

https://doi.org/10.5281/zenodo.8069920

Virtual BARTOC Editors Meeting
2023-06-23

Outline

- Introduction
- Numbers
- New Features
- Plans
- Challenges

Introduction

- Most comprehensive catalog of knowledge organization system (aka *terminologies*)
- Terminologies become more and more relevant!
- 3431 vocabularies and 122 registries
- 18 editors from 14 countries (+9 contributors)

BARTOC.org
Basic Register of Thesauri, Ontologies & Classifications

VZG

Technical Infrastructure

- Terminologies are stored in **JSKOS Server**: https://bartoc.org/api/

- Authentication with **ORCID**: no passwords are stored

- Web application written in JavaScript
  - Static pages in **Markdown**
  - Backend server in **NodeJS**
  - User Interface with with **EJS** (server) and **Vue** (client): needs cleanup

- Records contain information about **terminology APIs**

- Access to terminologies via this APIs using **cocoda-sdk** (API wrappers)

Everything is Open Source: https://github.com/gbv/bartoc.org

What happened in 2022/23?

- New terminologies and data updates
  - Large set of Korean KOS
    (see *Collecting Distributed KOSs and Building Meta Information* at NKOS 2022)
  - Information about APIs to access terminologies
- Switch from ILC1 to ILC2
- BARTOC Skosmos moved to VZG
- Improved Vocabulary Browsing
- Funding application for 3 years at German Research Foundation (DFG)
- New mailing list bartoc-editors@lists.gbv.de
- …

BARTOC.org
Basic Register of Thesauri, Ontologies & Classifications

Numbers

Vocabularies in BARTOC

- 3431 vocabularies, among them
    - ~950 classifications
    - ~800 thesauri
    - ~720 ontologies
    - ~350 glossary
    - ~200 name authority lists
- Most records created before 2020: 111 in 2021, **27** in 2022, **8** in 2023
- Records updated: 298 in 2020, 188 in 2021, **166** in 2022, **228** in 2023

BARTOC.org
Basic Register of Thesauri, Ontologies & Classifications

VZG

Terminology registries

Other catalogs and databases of terminologies

- 102 registries in 2021, 30 with API
- **122** registries in 2023, **52 with API**
  - 15 Skosmos
  - 9 SPARQL
  - 6 Ontoportal
  - 5 Ontology Lookup Service (OLS)
  - 3 JSKOS
- Actual number slightly lower, as defunct registries are not removed

Plan: query registries via API & merge data of their terminologies in BARTOC

8

Terminology API types

API	Vocabularies	Access implemented
Skosmos	276	yes
JSKOS API	98	yes
SPARQL	12	no (depends on)
Medical Subject Headings (SPARQL)	1	yes
Library of Congress	2	yes (partial)
Reconciliation API	4	yes
Skohub	5	yes
Medical Subject Headings SPARQL	1	yes (partial)
MyCore Classification XML file	3	yes
Library of Congress API	2	yes (partial)
Other Web Services	22	no (diverse)
OntoPortal	*many*	no
Ontology Lookup Service (OLS)	*many*	no

BARTOC.org

Basic Register of Thesauri, Ontologies & Classifications

VZGI

New Features

Vocabulary Browsing

▪ Via know APIs
▪ Usable but not beautiful

Hornbostel-Sachs-systematics

About   Access   Content   Identifiers

Search in vocabulary

Top concepts

Elektrophones

🔗 http://uri.gbv.de/terminology/hornbostelsachs/825670a8-676b-40a2-9367-de8d7694a708

Elektrophones^{en}
Elektrophone^{de}

📅 2017-08-18 | 📅 2017-08-18 | 📅 2018-11-29

↳ Analogue electronic instruments, modules and components
↳ Digital instruments, modules and components
↳ Software
↳ Hybrid analogue/digital configurations
↳ Electromechanical instruments and devices
↳ Electro-acoustic instruments and devices

BARTOC Skosmos

- Taken from Basel
- Still experimental
- Most KOS not included so far

BARTOC.org
Basic Register of Thesauri, Ontologies & Classifications

VZG

Plans

Funding proposal: main goals

- Improve **content and search** capabilities
  - Update and extend **information** about vocabularies
  - Redesign **search** with discovery and knowledge graph interface

14

Funding proposal: main goals

- Improve **content and search** capabilities
    - Update and extend **information** about vocabularies
    - Redesign **search** with discovery and knowledge graph interface

- **Integrate** with other systems of vocabulary management
    - **Import** from other terminology registries
    - Integrate **mappings and links** to indexed content

BARTOC.org
Basic Register of Thesauri, Ontologies & Classifications

VZG

Funding proposal: main goals

- Better **support** research data management
  - Support **editions and relations** between terminologies
  - Improve usability and workflows of **editing**
  - Update technical infrastructure

Funding proposal: main goals

- Better **support** research data management
  - Support **editions and relations** between terminologies
  - Improve usability and workflows of **editing**
  - Update technical infrastructure

- Support **terminology management**
  - Collection of tutorials, guidelines, software…
  - Information about hosting your own vocabularies

Challenges

Terminologies are not hyped

- *Knowledge Organization System* (KOS) not widely known
  beyond library and information science
- Other trends such as AI, LLM and full-text retrieval
  But: Knowledge Graphs are still a hot topic
- Lack of resources to properly document terminologies in BARTOC

Cultural Divide in terminology management

Communities with different language and domains of applications:

- **Vocabularies** in libraries, archives, museums ("humanities")
- **Ontologies** and terminologies in research data management ("sciences")

BARTOC is stronger in the first but known in both cultures.

Example: Mapping formats have been created independently with
JSKOS (https://gbv.github.io/jskos/jskos.html#concept-mappings) and
SSSOM (https://mapping-commons.github.io/sssom/) .
Reinventing wheels should be avoided.

Resources

- Daily updated BARTOC statistics at https://bartoc.org/stats
- BARTOC Issues https://github.com/gbv/bartoc.org/issues
- Editors mailing list bartoc-editors@lists.gbv.de
- Funding proposal https://doi.org/10.5281/zenodo.7673393 (German)

Don't hesitate to contact us!

https://bartoc.org/contact