

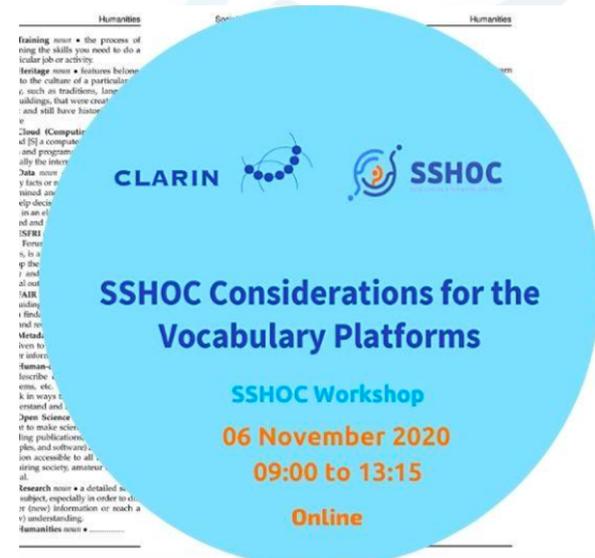
Choice of Vocabulary publication platforms for SSHOC

Monica Monachini

SSHOCWP3.1 leader

CNR-ILC

CLARIN - IT National Coordinator



The poster features a light blue circular background with the CLARIN and SSHOC logos at the top. Below the logos, the text reads: 'SSHOC Considerations for the Vocabulary Platforms', 'SSHOC Workshop', '06 November 2020', '09:00 to 13:15', and 'Online'. The background of the poster shows a blurred document with text related to 'Humanities' and 'Open Science'.

Humanities

Humanities

CLARIN

SSHOC

SSHOC Considerations for the Vocabulary Platforms

SSHOC Workshop

06 November 2020

09:00 to 13:15

Online

Humanities

Heritage

Heritage

Open Science

Open Science



SSHOC WP3

What

Lifting Technologies and Services into the SSH Cloud

How

- Contribute to the SSH Cloud development wrt both
 - infrastructure components and content, and
 - policies and governance
- Make CLARIN technologies also useful for the other SSH infrastructures



SSHOC T3.1

What

Provide background resources and tools to support better discovery of SSH research data in order to ensure better access and reusability.

How

- Data and services accessible and usable in SSH is a matter of providing
 - translated metadata in different languages and
 - a service to publish the vocabularies as a catalogue



Multilinguality

Searching, Browsing

SSHOC T3.1.3

What:

Milestone MS08:

Provide motivated recommendations for a vocabulary publication platform for SSHOC

How:

- Survey; interviews; main technological features to provide a SSHOC solution for the publication platform
- <https://drive.google.com/drive/u/0/folders/1XvRSb7hevHYWBemPpKEqAQDb9C8tsY5p>



Research and Innovation Action

Social Sciences & Humanities Open Cloud

Project Number: 823782

Start Date of Project: 01/01/2019

Duration: 40 months

Report on **Milestone MS8.0**

Choice of Vocabulary Publication platform for SSHOC

Dissemination Level	PU
Due Date of Milestone	30/06/2019, M8
Actual Achievement Date	30/06/2019
Lead Beneficiary/LTP	16. CNR
Work Package	WP3 Lifting Technologies and Services into the SSH Cloud
Task	3.1. Multilingual Terminology
Version	V1.1
Number of Pages	p. 1 – p. 18

Abstract: The goal of the activities in WP3.1.3 is to permit metadata-based discovery in different languages and one of

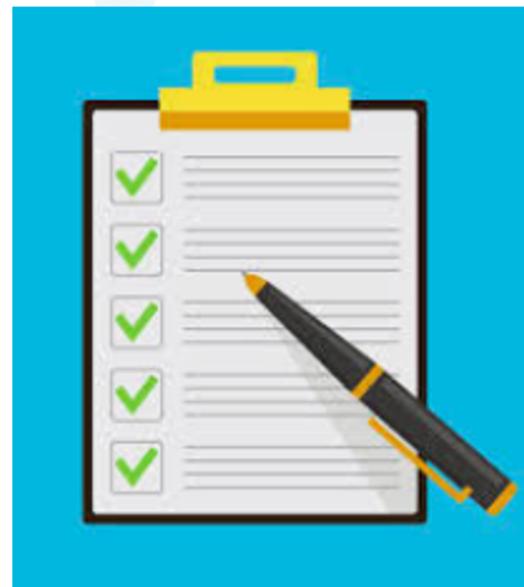


T3.1.3 The Survey of vocabulary management tools

Survey of existing systems for managing vocabularies to describe and allow discovery of research data in SSH infrastructures:

https://docs.google.com/spreadsheets/d/1cqeMujo0SwQHRogkuv2Xw6QV_PXIEGH-z74CwfWLT08/edit#gid=0

Matrix of systems/features



T3.1.3 The Matrix

Data model	Notes	Concepts	Synonyms	Hierarchy	Definitions or use notes	Concept schemes	Multilinguality translation	Import Export	Browser	Edit or	User management	Documentation	Linking Mapping Alignment	Examples	Versioning and history
------------	-------	----------	----------	-----------	--------------------------	-----------------	-----------------------------	---------------	---------	---------	-----------------	---------------	---------------------------	----------	------------------------

ACDH Vocabulary Service	https://vocabs.dariah.eu/en/about collaborative creation, maintenance and publication of vocabularies and taxonomies of any kind
THEMAS	https://github.com/isl/THEMAS/ (Thesaurus Management System) is a Web based system for creating and managing multi - faceted multilingual thesauri
CESSDA Vocabulary Service	https://vocabularies.cessda.eu/ discover, browse, edit and download controlled vocabularies in a variety of languages.
OpenSKOS	https://github.com/OpenSKOS/OpenSKOS web service based approach to publication, management and use of vocabulary data
Skosmos vocabulary browser	http://skosmos.org/ Open source web-based SKOS browser and publishing tool
iQvoc	https://iqvoc.net/ A Vocabulary Management System for the Semantic Web.
TemaTres	https://www.vocabularyserver.com/index.html open source vocabulary web application to manage and exploit vocabularies, thesauri, taxonomies
VocBench3	http://vocbench.uniroma2.it/ management of publication workflows, history, validation

T3.1.3 The Interviews

Does the vocabulary need a) an Editor platform for maintaining it, b) a browsing and publication platform to provide access to it?

What is the vocabulary used for?

Who uses it?

Who can browse/see it?

What is important in browsing, how would the user need to see the vocabulary?

Approximate number of terms/concepts

Who can edit the vocabulary, i.e. all interested parties or should editing rights be restricted to certain persons/roles?

Does it have hierarchy, that is, **broader and narrower terms**? (Yes/No). Does it have **related terms**? (Yes/No)

Do terms have **synonyms**/alternative labels? Does it have **definitions**?

Does the vocabulary exist in **different languages**? Please specify how many languages approximately.

Does the vocabulary have a **license**?

If translated, specify what content/elements are translated (terms, definitions, etc.). Who does the translation?

Is import of current vocabularies needed? If yes, what is the format of the vocabulary at present.

What export formats are needed?

Is API needed?

Is versioning needed, i.e. does the vocabulary have version number and publication date, or is it updated continuously?

Do users need to know what has been changed from last publication?

T3.1.3 Technical requirements

Standalone installation	Interoperability Protocols	Linked Data	API	Source Code	License
-------------------------	----------------------------	-------------	-----	-------------	---------

List of features that the SSHOC publishing platform should comply with as concerns interoperability protocols, provision of source code, API, the type of format and accessing the vocabularies.

https://docs.google.com/spreadsheets/d/1s5_StggMB1AburKRG4U6kK6oj5Dsc4weaGtXw0w1YNA/edit#gid=0

Vocabulary platform: SSHOC considerations

For the vocabulary publication platform in SSHOC

- not only a publishing tool, but a specialized tool for searching and browsing SKOS vocabularies
- not only a point of access, but also *editing* functionalities
- *linking/mapping/alignment* functionalities to reconcile vocabularies
 - Semi-automated alignment (suggestions, etc...)
 - Synchronization
 - Import/export
 - Role management

SSHOC Requirements for a vocabulary hosting platform

Data model	- SKOS I/O t(he most used data model in all vocabularies)
Structure	- support hierarchies of concepts may be an added value
Multilinguality	- Multilingual support is required to improve discovery and maximize the accessibility
Concept schemes:	- CS support is important in view of facilitating faceted search
APIs support	- important feature for vocabulary access



SSHOC Requirements for a vocabulary hosting platform

Browsing interface	<ul style="list-style-type: none">- browsing is regarded as an important added value if accompanied by specialized features
Advanced user management roles	<ul style="list-style-type: none">- system of roles for the maintenance and the incrementation of vocabularies- given the necessity for various partners to collaborate in different capacities- producing translations would optimally be user-friendly:
Versioning and status	<ul style="list-style-type: none">- Strict versioning with possibility to go back to previous states of the vocabulary and mappings and reconstruct the history of modifications- Status and contributions should be monitored; states. e.g., candidate, approved (wrt. the role of the contributors); or draft, initial review, final review, published.

SSHOC Requirements for a vocabulary hosting platform

- Import and export of thesauri in SKOS / RDF format;
- Unified access to all vocabularies
- Editing with collaborative functionalities
- Alignment functions between vocabularies and external resource
- Management of terms (status, language, URI, hierarchical structure --narrower, broader--, association (related terms), semantic relationships (synonyms), translations, facets; set of concepts (micro-thesauri);
- Management of different roles and workflows
- Management of versioning
- API services
- Friendly and intuitive interface, suitable for use by non-expert users
- Flexible to be adapted to new needs and standards

What's next



Discuss a more agreed-on solution for the vocabulary platform with experts

- further discuss hosting vocabulary platforms also investigating other possible solutions -- [Sept.2020](#)
- discuss/integrate the features -- [Nov..2020](#)