

Social Science and Humanities Open Cloud: why vocabularies matter

Iulianna van der Lek-Ciudin and Monica Monachini

CLARIN BAZAAR 2020

CLARIN IN SSHOC



Project:



Horizon 2020
European Union Funding
for Research & Innovation

Type of action & funding:
Research and Innovation action
(INFRAEOSC-04-2018)

Partners: 47

(20 beneficiaries + 27 LTPs)

SSH ESFRI Landmarks and Projects
& international SSH data infrastructures

Project budget:

€ 14,455,594.08

Duration: 40 months

(January 2019 – 30 April 2022)

Project website:
www.SSHOpenCloud.eu



Objectives:

- creating the social sciences and humanities (**SSH**) part of European Open Science Cloud (**EOSC**)
- maximising **re-use** through **Open Science** and **FAIR** principles (standards, common catalogue, access control, semantic techniques, training)
- interconnecting existing and new infrastructures (**clustered cloud infrastructure**)
- establishing appropriate **governance model** for SSH-EOSC

Why?

- Diversity of research methods and work practices in SSH
- Vocabularies are used to describe and access research contents.
 - Need to reconcile the current practices across domains
- CLARIN initiative to collect, register and harmonize SSH vocabularies/terminologies/ontologies to allow unified access to research contents.

Social Sciences plural noun • the study of society and the way people live; the subjects connected with this, for example history, economics, etc.

Humanities plural noun • literature, language, history, philosophy, and other subjects that are not a science, or the study of these subjects.

Open Science noun • the movement to make scientific research (including publications, data, physical samples, and software) and its dissemination accessible to all levels of an inquiring society, amateur or professional.

Cloud (Computing) noun • the cloud [S] a computer network where files and programs can be stored, especially the internet.

Data noun • information, especially facts or numbers, collected to be examined and considered and used to help decision-making, or information in an electronic form that can be stored and used by a computer.

Research noun • a detailed study of a subject, especially in order to discover (new) information or reach a (new) understanding.

SSH Open Marketplace noun • the only fully-integrated discovery portal which pools and harmonises the tools and services useful for the SSH research communities, offering a high quality and contextualised answer at every step of the research data

Library noun • a building, room, or organization that has a collection, especially of books, for people to read or borrow.

EOSC noun • an initiative aiming to build the European Open Science Cloud, a trusted and open distributed system for the scientific community, providing seamless access to data and interoperable services that address the whole research data cycle, from discovery and mining to storage, management, analysis and reuse across borders and scientific disciplines.

Training noun • the process of learning the skills you need to do a particular job or activity.

Heritage noun • features belonging to the culture of a particular society, such as traditions, languages, or buildings, that were created in the past and still have historical importance.



FAIR Data Principles noun • a set of guiding principles in order to make data findable, accessible, interoperable and reusable.

Metadata noun • information that is given to describe or help you use other information.

Human-centered adjective • used to describe computers, technology, systems, etc. that are designed to learn in ways that people can easily understand and learn.

Aimed at discovering which vocabularies are the most commonly used in research practice, define a methodology to map them, and align as far as possible with the SSHOC Reference Ontology.

Research data lifecycle noun • a model that describes and identifies the steps to be taken at the different stages of the research cycle to ensure successful data curation and preservation.

ESFRI noun • the European Strategy Forum on Research Infrastructures, is a strategic instrument to develop the scientific integration of Europe and to strengthen its international outreach.

Community noun • a group of people who have similar interests or who want to achieve something together.

SURVEY ON VOCABULARIES IN RESEARCH PRACTICE

Linguistics noun • the scientific study of the structure and development of language in general or of particular languages.

Tool noun • something that helps you to do a particular activity.

e-infrastructure noun • short term for Electronic Infrastructure, i.e., all ICT based resources (distributed networks, computers, storage devices, software etc.) and support operations which facilitate the collaboration among research communities by sharing resources, analysis tools and data.

Archaeology noun • the study of the buildings, graves, tools, and other objects that belonged to people who

How?

- Collect the SSHOC **requirements** for the vocabulary registry
- Organize online **information sessions** about vocabulary hosting platforms
- Organize a **SSHOC workshop** to discuss the results of MS08 survey: vocabularies, metadata, governance & sustainability
- Re-evaluate the MS08 vocabulary platform milestone wrt **new requirements**
- Select a **SSHOC vocabulary hosting platform**
- Create an **inventory of “known” vocabularies** with their metadata & **harmonize the metadata**
- **Deposit the metadata** in the SSHOC vocabulary hosting platform



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. 10

SSHOC Online Information Sessions: Open-Source Vocabulary Hosting and Management Platforms

Thursday, 3 September, 2020 - 15:00 to Wednesday, 30 September, 2020 - 16:30

As part of the SSHOC activities, CLARIN is organizing a series of online events to raise awareness about open-source vocabulary hosting and management platforms:

- 3 September 2020: First session - Introduction to Wikibase Vocabulary Platform
- 14 September 2020: Second session - Introduction to Skosmos Vocabulary Platform
- 30 September 2020: Third session - Introduction to Cessda Vocabulary Service

<https://www.clarin.eu/event/2020/sshoc-sessions-open-source-vocabulary-hosting-and-management-platforms>

SSHOC Considerations for Vocabulary platforms

- Save the Date: **Friday, 6 November, 9:00 – 12:30**
- In-depth discussion of the *MS8 SSHOC recommendations for vocabulary platforms* by Monica Monachini
- SSHOC WP updates on vocabularies
 - Vocabulary survey
 - Vocabulary mapping tool for Archaeology
 - How Dataverse supports external vocabularies
- Review of the online information sessions on open-source vocabulary platforms
- CESSDA ad CLARIN: How do they meet the SSHOC requirements for vocabulary platforms.
- Governance & Sustainability
- Discussion panel

Whom?

- Both academic and industry professionals who operate, integrate and manage vocabulary platforms, such as librarians, information specialists, knowledge management engineers, etc.

CLARIN in SSHOC: *Please visit also*

- **CLARIN services for the SSH**
Daan Broeder, Nicolas Larousse, Willem Elbers, Emanuel Dima
- **Facilitating Annotation of Collective Bargaining Agreements with Keywords Extraction**
Daniela Ceccon
- **The SSH Open Marketplace: it's relevance for the CLARIN communities and CLARIN services**
SSHOC: Stefan Buddenbohm (for SSH Open Marketplace), Laure Barbot (for DARIAH communities), Daan Broeder (for CLARIN communities)

Join our community!



www.sshopencloud.eu



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



[/in/sshopencloud](https://in.linkedin.com/company/sshopencloud)



info@sshopencloud.eu

