

SSH Conversion Hub

Mari Kleemola

Tampere University, Finnish Social Science Data Archive and
CESSDA ERIC

EOSC Symposium
Session: EOSC Interoperability
20 September 2023
Madrid, Spain



SSH Conversion Hub: product of the SSHOC project

- SSHOC project lifespan: Jan 2019 to April 2022
- Post-project, the SSH Research Infrastructures are organised as SSH Open Cluster
- SSH Conversion Hub was Part of *WP3 Lifting Technologies and Services into the SSH Cloud*



SSHOC, "Social Sciences and Humanities Open Cloud", received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-04-2018, grant agreement #823782.

SSH Conversion Hub Use Case

- Research team has found (meta)data that is in a format they are not familiar with (e.g., from other discipline). They want to integrate the (meta)data with their own data and with their analysis workflows. Optimally, without (too much) manual work.

Starting points for building the Conversion Hub

- There is a huge number of (meta)data formats and controlled vocabularies in use within SSH only
- There is an abundance of mappings, many in spreadsheet tables
- But not so many readily available conversion services or solutions
- We identified different kinds of existing conversion solutions:
 - Ready-to-use web-services
 - Installable software
 - Complex solutions that can entail the use of several services and manual steps ('conversion recipes')
- Researchers would benefit from (automated) conversion paths

Saved to this PC

Conversion Hub provides an inventory of solutions for conversions between data and file formats.

Use the search bar to discover solutions or use the facets to identify your area of interest.



Solutions

Expected Level of Knowledge

- advanced (16)
- expert (14)
- novice (7)

Community

- CLARIN (18)
- cross-community (13)
- CEESDA (8)
- DARIAH (5)
- ERIHS (2)
- LIBER (1)

Input format

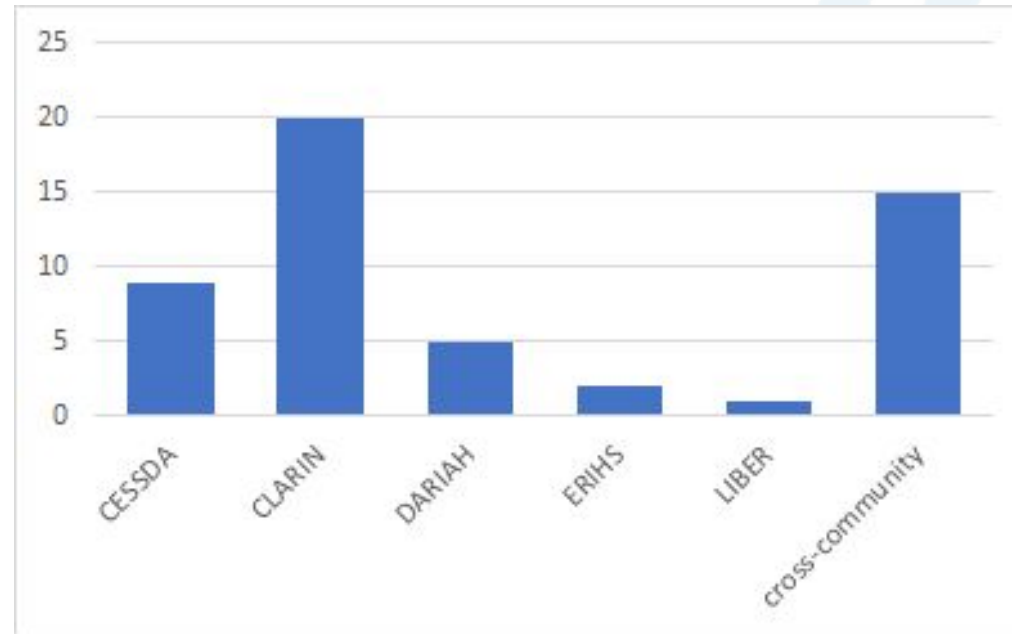
Displaying 1 - 41 of 41

Solution			
Title	Input format	Output format	Invocation type
3M	XML	RDF Triple	source code, web application
Audacity	WAV, AIFF, MP3, AU, FLAC, Ogg Vorbis	WAV, AIFF, MP3, AU, FLAC, Ogg Vorbis	local application
Beautiful Soup	HTML, XML	TSV, TXT, CSV, JSON	library
CHAMD	CHAT	PaQu metadata	source code
CMDI2DC	CMDI	Dublin Core	RESTful webservice
CMDI2Marc21	CMDI	MARC21 (XML)	RESTful webservice
CMDI2RDF	CMDI	RDF/XML	source code

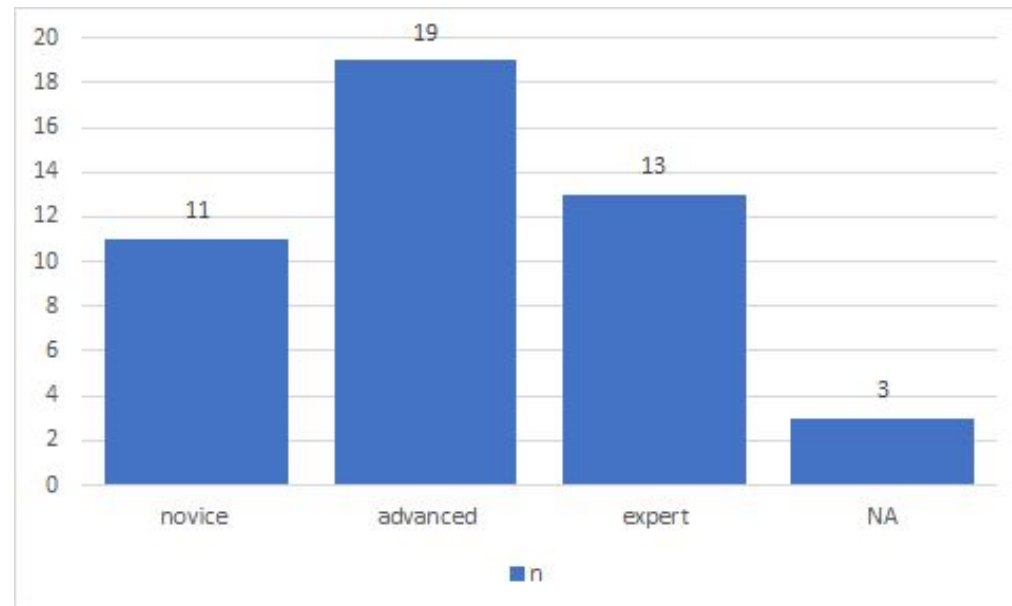
<https://conversion-hub.sshopencloud.eu/>

Number of solutions in the Conversion Hub by community

(the same solution can be listed under more than one community)



Expected level of knowledge needed to use the solutions



Interoperability of the SSH Conversion Hub itself

- Conversion Hub data model
 - Each solution in the Hub is described using several attributes
 - Many attributes use controlled vocabularies
 - Attributes were taken from, or mapped to, schema.org and SSHOCro
 - Structured metadata exposed via API
- Interoperable with the SSH Open Marketplace
<https://marketplace.sshopencloud.eu/>
 - Data models compatible, same vocabularies used
 - Conversion Hub is a source for SSHOMP

Findings

- Each SSH community has its own established traditions and customs that may cause interoperability issues with other communities.
- Metadata mappings/crosswalks typically 'study level' (basic findability), domain specific formats fine-grained
 - The existing diversity needs to be accepted since in most cases, the different formats and practices make good sense.
- For some solutions, adding documentation would be sufficient to enable use without extensive prior knowledge
- Some solutions are candidates for developing into easy-to-use web services
 - The licenses and ownership of solutions should be clearly stated

Some thoughts

- There have been plans to integrate ConversionHub into the SSHOMP for sustainability reasons
- SSH stakeholders have many more interoperability solutions within e.g. the RIs
- Originally it was thought that a suitable integration between the Conversion Hub and the Switchboard* can be achieved at the level of user interface, or at least at the level of exchange information about the suitable conversion services.
 - However, as the SSHOC project developed, it turned out to be better to use the SSHOMP as the rallying point of any information exchange.

* CLARIN service that helps researchers find tools that can process their data and that can be invoked from an external resource

Further information

- Conversion Hub: <https://conversion-hub.sshopencloud.eu/>
- Information on SSHOC website: <https://www.sshopencloud.eu/ssh-conversion-hub>
- Mari Kleemola, Katja Moilanen, Daan Broeder, Matej Ďurčo, Klaus Illmayer, Maurizio Sanesi, Emiliano Degl'Innocenti, Hervé L'Hours, Benjamin Mathers, Johan Fihn Marberg, Eleni Tsoulouha, Athina Kritsotaki, & Cesare Concordia. (2021). D3.6 Report on SSHOC format interoperability solution services, including new software.
<https://doi.org/10.5281/zenodo.5561604>



Thank you!