



Contribution ID: 219

Type: **Presentation**

FAIR mappings for data transformation and semantic alignment using Metadata Schema and Crosswalk Registry - Case Research Data Cloud (NII)

Wednesday 15 October 2025 14:44 (11 minutes)

Mappings are an essential component in making research data interoperable across infrastructures, domains and disciplines. Correspondences between official and de facto standards related to conceptual models, structures and vocabularies are required to share meaning and transfer information between both humans and machines. Despite their importance, these correspondences, mappings or crosswalks (i.e. collections of mappings [1]), can be hard to find or reuse as they are often scattered across different systems and available in a variety of formats and structures. Hence, making mappings FAIR (for Findable, Accessible, Interoperable, and Reusable) and machine-actionable has a great potential for reducing resources needed to create and maintain mappings. With FAIR principles the creation of high quality mappings can be supported by tools that help with discovery, access and evaluation of existing mappings at scale. Making mappings FAIR comes with specific requirements for mapping models and formats as well as associated metadata and services [1]. It has been proposed in [2] that these responsibilities should be handled by specialised mapping repositories to ensure the findability, availability, and reusability of these potentially highly valuable resources.

Metadata Schema and Crosswalk Registry [9] (MSCR) is a new and innovative service developed as part of the EU funded FAIRCORE4EOSC project and now operated by the IT Centre for Science, Finland. The MSCR offers multiple features that support and guide users through the different steps of the schema registration, mapping creation and development as well as publication. The MSCR supports versioning, lifecycle and content management and provides a persistent identifier for all published schemas, crosswalks and mappings for reliable referencing. By using MSCR, users have the possibility to map between any two registered schemas on the platform, which makes it possible to bridge ontologies and vocabularies as well as create crosswalks for data structure, format, and value transformation. The crosswalk editor supports mappings with different source and target element cardinalities (e.g. 1-to-1 and 1-to-many) and provides a set of processing functions that can be used to apply more complex mapping logic such as filtering or concatenation. Crosswalks created with the editor can be exported in different formats, such as SSSOM [7], XSLT and RML [8] to facilitate data transformation in other systems.

National Institute of Informatics (NII) in Japan manages the Research Data Cloud (RDC) [3], a research data infrastructure supporting research activities throughout their research data lifecycle. RDC consists of GakuNin RDM (GRDM) for research data management, JAICO Cloud for papers and data publication, and CiNii Research for discovery. The NII RDC application profile (NII RDC-AP) was recently developed [4] to enhance the interoperability of information exchanged among the services in the NII RDC. The challenge for NII is now to make the NII RDC-AP resources interoperable with other research data related application profiles in order to make research data within the NII RDC reusable. We present three mapping scenarios that showcase features of the MSCR and work towards concrete solutions to the interoperability challenges at NII and RDC:

1. **Transformation of metadata** from the GakuNin RDM to the NII data governance function for monitoring and supporting data management plans. There already exists a working solution called NII-DG-manager that implements JSON-to-JSON data transformation based on a custom mapping configuration. This scenario evaluates MSCR's mapping capabilities against the original python implementation from technical and user experience perspectives.
2. **Semantic alignment**; mapping of NII RDC-AP with other research data management platforms through ontology mapping between the NII-RDM ontology [5] and the SKG-IF ontology [6].
3. **Operationalization**; by creating a crosswalk between the mappings models of the MSCR and NII-DG-manager, which allows MSCR crosswalks (with certain restrictions) to be executed using NII-DG-

manager. This is an interesting meta-mapping type of case that can be extended to other target schemas as well.

The offering of the presentation is two-fold: First, we introduce an innovative and operational MSCR service, which supports a variety of mapping cases and fosters FAIR practices. Then, we present a practical example of the use and operationalisation of the MSCR in the context of NII national data infrastructure. The presentation targets a wide audience, including both newcomers and knowledge management practitioners (ontologists, metadata experts,...) who are interested in interoperable mapping practices.

References:

- [1] Le Franc, Y., Grau, N., Juty, N., Mejias, G., Reed, P., Ramezani, P., Poveda-Villalon, M., Garijo, D., Goble, C., & van Horik, R. (2025). D4.5 - Guidelines and methodology to create, document and share mappings and crosswalks (V1.1). Zenodo. <https://doi.org/10.5281/zenodo.15111167>
- [2] Broeder, D., Budroni, P., Degl'Innocenti, E., Le Franc, Y., Hugo, W., Jeffery, K., Weiland, C., Wittenburg, P., & Zwolf, C. M. (2021). SEMAF: A Proposal for a Flexible Semantic Mapping Framework (1.0). Zenodo. <https://doi.org/10.5281/zenodo.4651421>
- [3] Research Center for Open Science and Data Platform (RCOS), National Institute of Informatics. "Overview of the NII Research Data Cloud," <https://rcos.nii.ac.jp/en/service/> (accessed 2025-04-25)
- [4] Minamiyama, Y., Hayashi, M., Fujiwara, I., Onami, J.-ichi, Yokoyama, S., Komiyama, Y., & Yamaji, K. (2023). Toward the Development of NII RDC Application Profile Using Ontology Technology. Proceedings of the Conference on Research Data Infrastructure, 1. <https://doi.org/10.52825/cordi.v1i.260>
- [5] RDM Ontology Usage Guidelines, Research Center for Open Science and data platform (RCOS) in National Institute of Informatics, <https://rcosdp.github.io/RDM/> (accessed 2025-04-25)
- [6] The SKG-IF Ontology, Scientific Knowledge Graphs –Interoperability Framework (SKG-IF) WG <https://skg-if.github.io/data-model/ontology/current/skg-o.html> (accessed 2025-04-25)
- [7] <https://mapping-commons.github.io/sssom/>
- [8] <https://rml.io/>
- [9] Kesäniemi, J., Suominen, T., Broeder, D., & Puuska, H. Implementing interoperability - Metadata Schema and Crosswalk Registry approach to FAIR metadata mappings. EPiC Series in Computing

Primary authors: KESÄNIEMI, Joonas (CSC); Dr HIRAKI, Toshiyuki (National Institute of Informatics, the Research Center for Open Science and Data Platform)

Presenters: KESÄNIEMI, Joonas (CSC); Dr HIRAKI, Toshiyuki (National Institute of Informatics, the Research Center for Open Science and Data Platform)

Session Classification: Presentations Session 7: Open research through Interconnected, Interoperable, and Interdisciplinary Data

Track Classification: SciDataCon2025 Specific Themes: Open research through Interconnected, Interoperable, and Interdisciplinary Data