EPiC Series in Computing

Volume 105, 2025, Pages 58–68

Proceedings of EUNIS 2024 annual congress in Athens

Implementing interoperability - Metadata Schema and Crosswalk Registry approach to FAIR metadata mappings

Joonas Kesäniemi[1], Tommi Suominen[1], Daan Broeder[2] and Hanna-Mari Puuska[1]

[1] CSC – IT Center for Science, Espoo, Finland

[2] CLARIN ERIC, Utrecht, Netherlands

joonas.kesaniemi@csc.fi, tommi.suominen@csc.fi, d.g.broeder@uu.nl, hanna-mari.puuska@csc.fi

Abstract

Technical interoperability between information systems is usually thought to be solved in a dedicated manner by linking two or more technical systems together. However, in the context of research data and services management, the interoperability challenge can and should be more broadly scoped as that of interoperability between different descriptive systems. This leads to a need for flexible mapping between sets of descriptors and their values such as we see with the practice of metadata cross-walking.

In this paper, we are presenting a Metadata Schema and Crosswalk Registry (MSCR) implementation that can play an important part in implementing interoperability. The development of the MSCR is taking place in the context of the European Open Science Cloud (EOSC) and it aims to provide concrete solutions to information heterogeneity on semantic, structural, and syntactic levels. MSCR can be used to create a FAIR registry of schemas and crosswalks, which together provide the basis for validating and converting metadata documents in different formats.

Although the MSCR is being developed within the research data domain, applying its features to other domains, such as the exchange of educational information is straightforward. Higher education institutions (HEIs) could use the MSCR to not only better communicate the already existing work towards interoperability such as mapping documentation between standards for exchanging learning and research-related information, but also to enhance the concrete data integration infrastructure through the use of MSCR-managed schemas and crosswalks.

1 Introduction

Higher education institutions (HEIs), Research Institutes, and national and EU research infrastructures all exchange information widely between systems within their organizations but also extensively with each other as well as with other stakeholders. For example, national and international mobility of students and researchers is easier when their personal data and information on their research and education activities can be exchanged automatically. HEIs also report information on education, research, and administrative functions to, for example, national, EU level and other authorities and funders, which in an ideal situation would be handled in an automated way.

Still, on many occasions, learners, teachers, and researchers input their own information in various services and sometimes the information is provided even by paper. Advanced integration between systems promises that the administrative burden can be reduced as the data only needs to be entered once, the data exchange becomes faster, and the quality stays higher as manual tasks are reduced and replaced with automated processing. Interoperable information systems can exchange data with one another, whereas semantic interoperability makes it possible to make sure that the exchanges make sense (Heiler, 1995). The same sentiment is summarized by the European Interoperability Framework (EIF) as "what is sent is what is understood" (Commission & for Digital Services, 2017). The EIF model further divides interoperability into four layers: legal, organizational, semantic, and technical. These levels are also covered by the principles put forward in the upcoming Higher Education Interoperability Manifesto (Commission, 2023)

At a technical level, the integration of information systems is typically implemented by utilizing application programming interfaces (APIs), either for reading or writing data or metadata. Some form of transformation or mapping of the data is often required to make the data provided by an API compatible with the information model of the target system. Creating a single crosswalk between data models of two systems might not be a big issue, but it can lead to exponential expansion of integrations. With the point-to-point integration approach, n systems needing to integrate (integrations are directed i.e. an integration from a to b does not enable an integration from b to a) leads to a need of n*(n-1)*2 integrations.

The use of a core set to address the scaling problem posed by having to transform between many different schemas is a well-known and established strategy. Also, as the SEMAF report (Broeder, et al., 2021) points out, crosswalks don't always have to be "complete" i.e. cover all the schema elements, but rather be purpose-built for a specific task and extendable as needed. This requires of course good provenance tracking and sharing options as part of crosswalk maintenance. Similarly to the core-set approach identifying a common information model (or ontology) that maximizes coverage, and that all services create crosswalks to and from is considered helpful by many. This enables reading in and exporting data with translation to only one external data model. Once you can ingest data in this generic format, you can ingest data from any service that has created a crosswalk to it. However, if not careful, it may derail into a (non-scalable) attempt to create an all-encompassing ontology.

The solution we describe in this paper to the above problem is called the Metadata and Schema Crosswalk Registry (MSCR), which is being developed in the context of the FAIRCORE4EOSC project (FAIRCORE4EOSC — faircore4eosc.eu). The MSCR tackles the interoperability challenges stemming from information heterogeneity, for which the solutions can classified according to (Ouksel, 1999) as semantic, structural, and syntax interoperability. The MSCR allows registered users and communities to create, register, and version schemas and crosswalks. The published content can be searched, browsed, and downloaded without restrictions. The MSCR also provides a mechanism to facilitate the transformation of data from one schema to another via registered crosswalks.

One of the key objectives of the MSCR is to make the crosswalks/mappings more FAIR, which refers to the four now well-established principles of Findability, Accessibility, Interoperability, and Reusability (Wilkinson et al., 2016). The MSCR provides projects and individual researchers with the possibility to manage their metadata schema and/or relevant metadata schema crosswalks. The schema and crosswalks are shared with the community for reuse, and the extension is supported by a proper versioning mechanism. The fact that all schemas, mappings between parts of schemas, and the crosswalks composed from mappings are identified with persistent identifiers (PIDs), significantly improves the ability to reference MSCR content and makes them more findable and reusable. This ability to cite, point at, and reuse crosswalks means that a researcher can create their own version for a specific need, but thereby also creating the need to manage those versions i.e. tracking the provenance of the crosswalks. Since the crosswalks have authorship and PIDs, they become a citable research object in their own right as is proper given their role in research.

The development of the MSCR is taking place in the context of the European Open Science Cloud (EOSC). *"The ambition of the European Open Science Cloud, known as EOSC, is to develop a 'Web of FAIR Data and Services' for science in Europe. The implementation of EOSC is based on a long-term process of alignment and coordination pursued by the Commission since 2015." (*https://eosc.eu/eosc-about) However, this generic new service may have far-ranging impacts well beyond this initial domain of application.

2 Related work

The potential of using metadata schema registries to help with metadata interoperability has been recognized for decades now. Basic functional requirements for a metadata schema registry (e.g. schema mappings and version management) were presented in (Nagamori & Sugimoto, 2006) based on their work on schema registries since the late nineties. In their vision, a network of schema registries would provide schema-based services across different languages and communities.

Research Data Alliance (RDA) has also been active in the research data-related schema registry work for a long time. Metadata Standards Directory (Metadata Standards Catalog — rdamsc.bath.ac.uk) was published as an output of the Metadata Standards Catalog Working Group in 2015, which was based on an earlier work by the Digital Curation Center (DCC) on its Disciplinary Metadata Catalog (Disciplinary Metadata | DCC — dcc.ac.uk). The Metadata Standards Directory was eventually superseded by the RDA Metadata Standards Catalog (RDAMSC) operated by the University of Bath (Metadata Standards Catalog — rdamsc.bath.ac.uk). The metadata about schemas in the RDAMSC covers for example information about maintaining organizations, documentation, representations (e.g. XML Schemas), known users, tools, and fields of science. It also hosts a list of mappings between cataloged schemas with links to documentation and implementations bundled with information about responsible organizations and persons. One of the more recent efforts in schema cataloging with a clear emphasis on FAIR principles, is the FAIRSharing service* that includes not only standards but also databases, repositories, and data policies. At the time of writing, the service had more than 1,700 standards from which around 800 were categorized as "terminology artifact" and around 600 as "model and format". This makes for around 1,400 metadata records about potential mapping sources or targets for the MSCR.

When it comes to tools for creating and maintaining mappings, two distinct paths can be recognized. First, there are tools and services with a vocabulary or ontology-related background that focus on semantic mappings between concepts. Second, are the tools that support structural and syntactic mappings between schemas expressed in various schema definition languages (e.g. XML Schema). Examples of the first trend are OntoPortal (Jonquet, ym., 2023), Cocoda (Cocoda — coli-conc.gbv.de), and Vocabulary matching tool (Ceri Binding). Cocoda and Vocabulary matching tool are standalone tools for creating and maintaining mappings, whereas the OntoPortal is primarily an ontology repository with additional support for automatically created and manually curated mappings. The Data Modeling Environment (DME)† is a good example of the latter group. With the DME users can create and maintain vocabularies, schemas, and mappings and it even provides an API for running the data transformations with instance data. Particularly the mappings features of DME are very sophisticated

* https://fairsharing.org/
† https://dme.de.dariah.eu/dme/registry/

(Steyer & Gradl, 2019). The functionalities where DME is lacking are related to core content management aspects such as rich content metadata, provenance, and versioning.

OxO (Jupp, ym., 2017) is an example of a pure mapping repository. It brings together ontology mappings from multiple sources and harmonizes their identifiers to provide an integrated graph for users to explore direct and indirect mappings between ontology terms. The goal of the service is to provide curators with a collaboration platform for improving the mappings and to collect background knowledge for automated ontology matching tools.

Finally, as the goal of the MSCR is to facilitate the conversion of metadata from one format into another, we must mention the numerous Extract-Transform-Load (ETL) systems and data brokers that have already been doing exactly that for years in production environments. The main difference is that when the ETL and similar systems are implementing data mappings for specific data sources, the MSCR is based on schema mappings for semantic, structural, and syntactic levels to support interoperability for any data source conforming to the schema.

3 Metadata Schema and Crosswalk Registry

The Metadata Schema and Crosswalk Registry is built on the work of the Finnish Digital and Population Data Services Agency on the Interoperability Platform (Alonen & Remes, 2016). The MSCR functionality is based on the Data Vocabularies tool of the Interoperability Platform suite of applications, which in addition to basic features such as authentication, authorization, and reusable user interface (UI) components also includes an API implementation for creating, searching, and modifying graph-based models expressed in RDF[‡]. By reusing and building on top of a well-maintained code base, the development team has been able to start building new crosswalk-related features quickly.

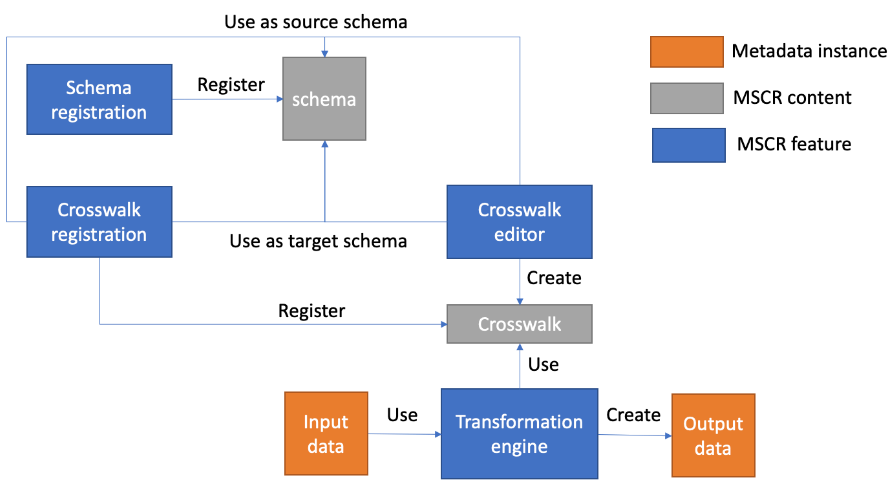

Figure 1: Crosswalk editor is used to create mappings between registered source and target schemas. Crosswalks can be operationalized through a transformation engine that can be used to convert input data conforming to the input schema to output data that adheres to the target schema.

‡ https://www.w3.org/TR/rdf11-concepts/

As the name suggests, schemas and crosswalks are the basic content types managed by the MSCR. There are two ways of adding content: registration of existing content and creation of new content from scratch (Figure 1). Existing content can be either uploaded as part of the registration process or referenced from an external source (e.g. document repository). With the first option, MSCR is considered the primary source of the content with hosting and maintenance responsibilities. The distinction between *hosted* and *referenced* content affects how some of the metadata for the content is handled. The MSCR mints a handle identifier for each piece of content. In the case of hosted content, the handle identifier is directly associated with both the metadata and the schema content, but in the case of the referenced content, the MSCR-assigned identifier points only to the metadata, and the original identifier serves as the source for the actual content (Figure 2). Continuing the use of already published existing identifiers is crucial for preserving the provenance of registered content as well as communicating responsibilities regarding content maintenance. Regardless of the origin of the content, a copy of the content will always be stored on the MSCR's side.

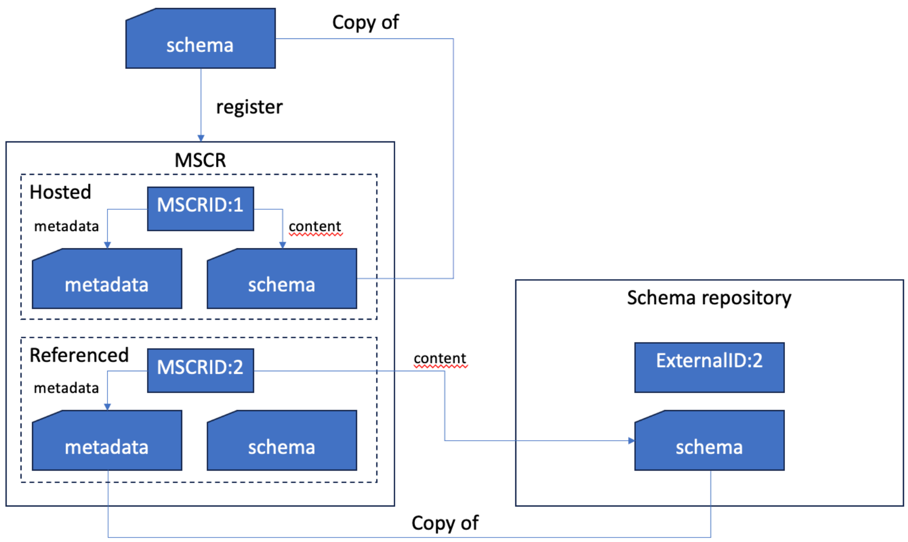

Figure 2: Differences in PID resolving with MSCR hosted and externally hosted (i.e. referenced) schema content.

There is currently no way of creating schemas from the ground up using the MSCR, so the users need to register existing schema content. In the context of the MSCR, we use the term schema quite loosely to refer to a document that defines the structure of information in a specific format. The goal is to support schemas expressed in one of the following schema specification languages (with format in parenthesis): CSV (delimited text file), JSON Schema (JSON), XML Schema (XML), SKOS[§], SHACL[**], RDFS[††] and OWL[‡‡]. In this context, RDF is used to denote any commonly supported RDF serialization. In addition to the previously listed formal schema definition languages, it is also possible to add a schema description as a PDF file. The reasoning behind supporting purely informative schema descriptions (PDF) is the fact that it still is a very common and useful piece of documentation that can

[§] https://www.w3.org/TR/skos-reference/
[**] https://www.w3.org/TR/shacl/
[††] https://www.w3.org/TR/rdf-schema/
[‡‡] https://www.w3.org/TR/owl2-overview/

be used to create a more structured version of the schema. We also include knowledge organization systems expressed as SKOS (in RDF) as part of the schema content, because they are common subjects for semantic mappings. These vocabulary-to-vocabulary types of mappings are also required for metadata instance-level transformations (e.g. mapping values between code lists as part of the data conversion).

Registered schemas are a prerequisite for adding crosswalk content as both registration and creation of crosswalks require explicit references to source and target schema (Figure 1). Many existing crosswalks are documented in PDFs or spreadsheets that have not been meant for automated processing. On the other hand, there is an abundance of scripts, code, and stylesheets that implement even the most complex transformations. These existing crosswalks can be registered to the MSCR either by uploading the content files or by providing a reference to externally maintained content, such as a GitHub commit or a Software Heritage Persistent Identifier (Di Cosmo;Gruenpeter;& Zacchiroli, 2018). The added value that MSCR can provide for these existing solutions is the references to a specific source and target schema descriptions that can be used for searching (e.g. get all crosswalks that work on DataCite 2.1[§§] and potentially for validating both input and output metadata. This provides a powerful way of making existing solutions more findable and reusable. For semantic mappings it is also possible to import a set of mappings using Simple Standard for Sharing Ontology Mappings (SSSOM) serialized as tab-separated values file (Matentzoglu, et al., 2022).

While registering existing crosswalks can make them more FAIR, the source of the true power of MSCR lies in the crosswalk editor that can be used to create new crosswalks and edit them. It should be noted that at this point the editor can only be used to edit crosswalks created with editor and registered SSSON-based content. It is not possible to edit for example an existing code-based crosswalk.

In the context of MSCR's crosswalk editor, a crosswalk is defined as a set of mappings between source and target schema structures. Mappings should be ideally descriptive enough to be able to fully define the transformation of metadata instances on a structural and syntactic level. The mapping model of the MSCR supports 1-to-1, 1-to-many, many-to-1, and many-to-many types of structural mappings with the possibility to modify the metadata instance values with a predefined set of operations (e.g. addPrefix or convertUnit). The same mapping model and functionality can be used to create mappings between any of the schema description formats supported. For example, one could map an XML schema element to a class in an ontology or a concept in a SKOS vocabulary to associate it with machine actionable semantics. In the case of two vocabularies, one could map the vocabularies together and merge them, creating a third vocabulary that can be used for example to aggregate metadata instance values expressed using the source vocabulary with the values from the target vocabulary (i.e. many-to-one mappings). The details of the mapping model are out of the scope of this paper, but Figure 3 provides an overview of different kinds of mappings and how they are used in the MSCR.

[§§] https://schema.datacite.org/archive/kernel-2.1/index.html

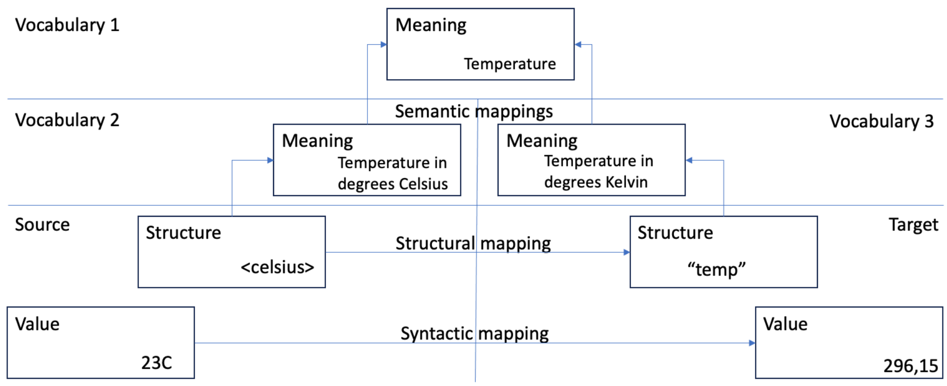

Figure 3: Example of data instance transformations on syntactic and structural levels with helpful semantic mappings. <celsius> XML element and "temp" JSON property are mapped to temperature measurement using specific units, which in turn are linked to a more generic temperature measurement. All of this information can be stored in the MSCR. Links between structural and ontological schemas (in the broad sense of the word) can be used to help with the mappings between XML and JSON schemas.

The MSCR views mappings primarily in the context defined by the crosswalk's metadata. For this reason, the individual mappings are not given their own handles, but they are identified by adding part identifiers[***] to the crosswalk's handle identifier. There are also no restrictions on how comprehensive a crosswalk should be. Adding just a few mappings might suffice in some cases; in others, there might be a need to map the whole target schema. Source and target schemas will help the user to find candidate crosswalks for a specific use case and the metadata of the crosswalk and included mappings should provide information for evaluating the suitability of each candidate crosswalk for the task at hand.

We argue that managing metadata about schemas, crosswalks, and mappings in a registry is already a worthwhile undertaking in itself since it increases the findability and reusability of important resources. However, schema and crosswalk registries can also be seen as the only means to an end, which is the transformation of one metadata instance to another. The MSCR will include features that support this kind of task in the form of a transformation engine component. This component could be used for example as part of the crosswalk editor to test changes with sample input metadata. More realistic transformation tasks could be implemented either by a managed transformation service or by local deployments of the containerized version of the component that would retrieve its configuration, i.e. crosswalks and schemas, from the MSCR.

In conclusion, the features of the MSCR are applicable in multiple domains and the sharing and caring for crosswalks - making them FAIR – is useful in its own right – but the most significant value of this service will be realized through the operationalized metadata transformations.

[***] https://epic.grnet.gr/docs/api-partial/

4 Potential impact for information exchange in higher education

Both in research and education, there is a huge amount of different standards in Europe. Application opportunities in other higher education fields could be e.g. supporting international student exchange and international degree students. This could potentially have considerable positive effects on e.g. European university alliances.

In the FAIRCORE4EOSC project, we are applying the MSCR to Current Research Information Systems (CRISes) where the most commonly adopted model is the Common European Research Information Format (CERIF) (Jörg, 2010). CERIF supports the exchange of information between CRISs and with other information systems, such as institutional repositories, research data archives, and others, both on a local, national, or international level (Jeffery;Houssos;Jörg;& Asserson, 2014) (Jörg, 2010). The MSCR can be used to create crosswalks between, for example, an institutional CRIS system and the EOSC Research Discovery Graph (RDGraph) that provides the linked information to be used across EOSC resources and communities (Suominen, 2022). All that is needed for any given service is the capacity to ingest data in its own native data model. No further integrations to other data models are needed.

The MSCR can also be applied in other HEI-related domains. For studies, degrees and credits, and learning opportunities as well as skills and qualifications data, ELMO and ELM are the basic standards in Europe (Fridell, ym., 2022). However, many institutions, countries, and systems apply their very own metadata models, which do not follow these standards. Also, there are currently very few crosswalks available for converting between models in different domains, like science and education, for example when a student becomes a researcher.

Some of the potential benefits of using MSCR in the HEI domain include:

- MSCR could be used to create a cross-domain, unified, and machine-actionable registry of standards for the HE environment, which would make a comparison of schemas easier and could maybe help even with some consolidation efforts in the form of core schemas for different domains.
- Turning existing spreadsheet mapping exercises into runnable mappings would increase the value of already completed work.
- Documenting or migrating code-based mappings to MSCR makes the crosswalks and their details accessible to a wider audience within the HE, which can help with getting the semantics right and in general error spotting.
- Evolution of schemas and crosswalks can be managed using the versioning features of MSCR. Identifiers and single sources of truth make it easier to discuss changes in schemas and crosswalks.
- In the higher education sector, MSCR can be used in e.g. creating conversion between the two data formats, ELMO and ELM, as is being pursued in the DC4EU project (EMREX, 2024) but

also more comprehensive efforts undertaken within the European Higher Education Interoperability Framework[†††]

5 Discussion

Building a new solution that integrates semantic, structural, and syntactic mappings with vocabularies, schemas, and ontologies comes with its technical complications. Still, more importantly, it involves some difficult user experience and training challenges that need to be addressed before MSCR can grow to its full potential.

Even with the consolidation of metadata models there will still be a need for crosswalks when for example cross-discipline or cross-continent interoperability of systems is required. To prevent the combinatorial explosion of mappings, the SEMAF report emphasizes the importance of efficiently managing and sharing all known relevant mappings and being able to reuse previous work (Broeder, et al., 2021). The use of MSCR might even promote a "mapping frenzy" at least initially, but we are discussing different options for steering MSCR users towards more efficient use of existing work, such as standard schemas, through technical and non-technical means.

In its current form, the MSCR can be populated with either existing content or by using the crosswalk editor to manually create new content. In this era of large language models (LLM), one interesting avenue for further development of the crosswalk editor would be to use LLMs to generate the first suggestion for the crosswalk that the user could then manually refine. As the curated content of MSCR grows, the content could also maybe be utilized as training material for new machine-learning models specializing in schema mapping. Another use case for LMMs would be the semi-automated extraction of structured schemas and crosswalks from content registered as PDFs or other formats meant primarily for human consumption.

To attract and sustain long-term interest from users in the MSCR it is essential to also provide a minimal level of editorial effort. Ideally, this would be provided by user communities that find the MSCR a cost-effective and hopefully long-term (commonly) maintained solution. However, experience tells us that commitment from one or more larger partners is indispensable to inspire sufficient trust and kick-start the whole process of finding appropriate modes for collaboration and governance solutions.

A unified and accessible repository of schemas and especially crosswalks can be considered an asset for any organization or community. Having the data in a machine-actionable format, e.g. schemas for validation and crosswalks for data transformation, makes the proposition even more appealing. If the HEIs choose to embrace the idea of a metadata schema and crosswalk registries, the first step should be the identification and registration of core schemas for different domains, or collaboratively use the MSCR to come up with one.

6 Acknowledgements

The work described in this paper has taken place in the context of the FAIRCORE4EOSC project. FAIRCORE4EOSC has received funding from the EU's Horizon Europe research and innovation programme under Grant Agreement no. 101057264.

[†††] https://education.ec.europa.eu/focus-topics/digital-education/digital-education-hub/workshops-and-working-groups/european-higher-education-interoperability?facets__permanent%7Cfield_eac_tags=526

References

Alonen, M., & Remes, S. (2016). Interoperability Workbench–Collaborative Tool for Publishing Core Vocabularies and Application Profiles. *International Conference on Dublin Core and Metadata Applications*, (pp. 90–93).

Broeder, D., Budroni, P., Degl'Innocenti, E., Le Franc, Y., Hugo, W., Jeffery, K., . . . Zwolf, C. M. (2021, March). SEMAF: A Proposal for a Flexible Semantic Mapping Framework. *SEMAF: A Proposal for a Flexible Semantic Mapping Framework*. Zenodo. doi:10.5281/zenodo.4651421

Carlisle, D. (2010, April). *graphicx: Enhanced support for graphics.* Retrieved from http://www.ctan.org/tex-archive/ help/Catalogue/entries/graphicx.html

Ceri Binding, c. (n.d.). Vocabulary Matching Tool — vmt.ariadne.d4science.org. *Vocabulary Matching Tool — vmt.ariadne.d4science.org*.

Cocoda — coli-conc.gbv.de. (n.d.). *Cocoda — coli-conc.gbv.de*.

Commission. (2023). Bridging Boundaries: A Manifesto for a European Higher Education Interoperability Framework - DRAFT. *Bridging Boundaries: A Manifesto for a European Higher Education Interoperability Framework - DRAFT*.

Commission, E., & for Digital Services, D.-G. (2017). *New European interoperability framework – Promoting seamless services and data flows for European public administrations.* Publications Office. doi:doi/10.2799/78681

Di Cosmo, R., Gruenpeter, M., & Zacchiroli, S. (2018, September). Identifiers for Digital Objects: the Case of Software Source Code Preservation. *iPRES 2018 - 15th International Conference on Digital Preservation*, (pp. 1-9). Boston. doi:10.17605/OSF.IO/KDE56

Disciplinary Metadata | DCC — dcc.ac.uk. (n.d.). *Disciplinary Metadata | DCC — dcc.ac.uk*.

EMREX. (2024, 04 15). *EMREX Newsletter*. Retrieved from https://emrex.eu/2023/09/06/emrex-newsletter-september-2023

FAIRCORE4EOSC — faircore4eosc.eu. (n.d.). *FAIRCORE4EOSC — faircore4eosc.eu*.

Fridell, T., Vangen, G., Mincer-Daszkiewicz, J., Norder, J. J., Kohtanen, J., Drvodelić, I., & Bacharach, G. (2022). Interoperability of educational data demands standards. *European Journal of Higher Education IT, 1*.

Heiler, S. (1995, June). Semantic interoperability. *ACM Comput. Surv., 27*, 271–273. doi:10.1145/210376.210392

Jeffery, K., Houssos, N., Jörg, B., & Asserson, A. (2014, February). Research information management: the CERIF approach. *Int. J. Metadata Semant. Ontologies, 9*, 5–14. doi:10.1504/IJMSO.2014.059142

Jonquet, C., Graybeal, J., Bouazzouni, S., Dorf, M., Fiore, N., Kechagioglou, X., . . . Musen, M. (2023). Ontology Repositories and Semantic Artefact Catalogues with the OntoPortal Technology. In T. R. Payne, V. Presutti, G. Qi, M. Poveda-Villalón, G. Stoilos, L. Hollink, . . . J. Li (Ed.), *The Semantic Web – ISWC 2023* (pp. 38–58). Cham: Springer Nature Switzerland.

Jupp, S., Liener, T., Sarntivijai, S., Vrousgou, O., Burdett, T., & Parkinson, H. E. (2017). OxO - A Gravy of Ontology Mapping Extracts. In M. Horridge, P. Lord, & J. D. Warrender (Ed.), *Proceedings of the 8th International Conference on Biomedical Ontology (ICBO 2017), Newcastle-upon-Tyne, United Kingdom, September 13th - 15th, 2017. 2137.* CEUR-WS.org. Retrieved from https://ceur-ws.org/Vol-2137/paper_{2}{7}{.}pdf

Jörg, B. (2010). CERIF: The common European research information format model. *Data Science Journal, 9*, CRIS24–CRIS31.

Matentzoglu, N., Balhoff, J. P., Bello, S. M., Bizon, C., Brush, M., Callahan, T. J., . . . others. (2022). A simple standard for sharing ontological mappings (SSSOM). *Database, 2022*.

Metadata Schema and Crosswalk Registry (MSCR) | FAIRCORE4EOSC — faircore4eosc.eu. (n.d.). *Metadata Schema and Crosswalk Registry (MSCR) | FAIRCORE4EOSC — faircore4eosc.eu*.

Metadata Standards Catalog — rdamsc.bath.ac.uk. (n.d.). *Metadata Standards Catalog — rdamsc.bath.ac.uk*.

Metadata Standards Directory. (n.d.). *Metadata Standards Directory*.

Nagamori, M., & Sugimoto, S. (2006). A metadata schema registry as a tool to enhance metadata interoperability. *TCDL Bulletin, 3*.

Ouksel, A. M. (1999). Semantic interoperability in global information systems. *ACM Sigmod Record*, 5-12.

Sansone, S., McQuilton, P., Rocca-Serra, P., Gonzalez-Beltran, A., Izzo, M., Lister, A., . . . Community, F. (2019). FAIRsharing as a community approach to standards, repositories and policies. *Nature Biotechnology, 37*, 358-367.

Steyer, T., & Gradl, T. (2019). A research-oriented and case-based data federation for the Humanities. doi:10.5281/zenodo.2536107

Suominen, T. (2022). Developing EOSC-Core components to enable a FAIR EOSC ecosystem. *Strategic Membership Meeting 2022 – Autumn* . Nijmegen: euroCRIS.

Tobias Gradl, U. o. (n.d.). DARIAH-DE Data Modeling Environment — dme.de.dariah.eu. *DARIAH-DE Data Modeling Environment — dme.de.dariah.eu*.

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., . . . others. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data, 3*.