

## European Integration of National-level services – a look on the case study of FAIRCORE4EOSC project

This paper presents the current work of FAIRCORE4EOSC-project's case study relating to European CRIS landscape and possible data exchanges within CRIS-domain. Also related is the implementation of RAiD identifier for research projects and related activities. The work has started in June of 2022 and continues for the length of FAIRCORE4EOSC's scheduled end in May 2025. Participating in the case study are the Finnish national CRIS, Research.fi, developed by CSC – IT Center for Science, concept study on Dutch Open Knowledge Base by SURF and the Greek HELIX service maintained by Athena RC.

### Background

Vast amounts of research information are constantly being collected and stored in various different locations e.g. field-specific services, international databases, universities' systems, libraries, scientific publishers and research funders' registers. Achieving the interoperability of research information via common solutions provides for many aspects of research and any related stakeholders. The main difficulty of standardization and interoperability of data at the European level is the variety of international, national/regional and organizational CRIS systems and their data models, which make for a situation where data exchange between these solutions is made difficult.

Main benefit of CRIS systems in general, be it implemented on any of the abovementioned levels, is the quality of metadata which is achieved by both the highly structured format in which the data on research outputs is collected and the extent in which the metadata is reported from. Well structured metadata enables sophisticated analysis on research and the extent provides for wider perspective.

### FAIRCORE4EOSC

The FAIRCORE4EOSC project<sup>1</sup> focuses on the development and realisation of core components for the European Open Science Cloud (EOSC). Supporting a FAIR EOSC and addressing gaps identified in the Strategic Research and Innovation Agenda (SRIA). Leveraging existing technologies and services, the project will develop nine new EOSC-Core components aimed to improve the discoverability and interoperability of an increased amount of research outputs. The components delivered within the project are also enabling the EOSC PID Infrastructure, an EOSC research software infrastructure, support for sharing and access to metadata schemas and crosswalks and offer advanced research-intent driven discovery services over all EOSC repositories.<sup>2</sup>

Five case studies will be implemented during the project to provide the development teams with user-driven requirements and to act as testbeds for the components. Each case study will use multiple components developed by the project to validate their functionalities and

---

<sup>1</sup> Core Components Supporting a FAIR EOSC, <https://doi.org/10.3030/101057264>

<sup>2</sup> <https://faircore4eosc.eu/>

provide user feedback. Case studies will also be instrumental to produce long-term benefits to their users, ensuring sustainability of the integration of the new components. The case studies will not be the only validation instrument adopted by the project.

## European Integration of National-Level Services

Case study of European Integration of National-level Services as part of FAIRCORE4EOSC aims to expand the use and the coverage of CRIS-like information on European level, which would be invaluable source of information for e.g. research assessment or to highlight EOSC-related contributions. The case study involves three national CRIS systems (Research.fi / Finland / CSC, Open Knowledge Base / Netherlands / SURF, Helix / Greece / ARC) in which the EOSC Core components developed within the FAIRCORE4EOSC will be piloted.

Case study is led by CSC and it tries to tackle few of crucial obstacles within European CRIS-domain, which are: 1) Inability to fully exchange data within the Europe on CRIS-like metadata<sup>3</sup>, 2) Assigning RAiDs within CRIS-domain for research projects and research related activities and 3) Operationalizing metadata data models / schemas of CRIS system and maintaining these crosswalks. Also a part of the case study's goals are to have a working template on enriching RDGraph of OpenAIRE with more research entities than are currently available in OpenAIRE and current integrations between CRISs and OpenAIRE.

As the interoperability of research information is highly central in this case study and resolving issues with the interoperability of CRIS-like metadata is prioritized, the Metadata Schema and Crosswalk Registry (MSCR) is identified as crucial component. Two primary requirements were identified: 1) Operationalize metadata mappings between various national CRIS services (e.g. [Research.fi](http://Research.fi)) and CERIF data model and 2) Ability to refer to usage of and the updates of these mappings, by individuals and especially by services and communities. This is achieved using CERIF data model (*Common European Research Information Format*) maintained by euroCRIS, which is a common for exchanging information between organizational and national CRIS systems & OpenAIRE. Work also support the current refactoring project of CERIF, maintained by euroCRIS.<sup>4</sup>

The aim is to realize the potential in combining CRIS-like information from national sources on European level, aggregating this information to RDGraph and OpenAIRE.

Research Activity Identifier Service i.e. the adoption of RAiD identifiers is also a key in the case study and how it related to CRIS-domain. In operationalizing RAiD in, two requirements were identified: 1) Take into use the European RAiD Registration Agency for national CRIS systems and 2) Managing and upkeep of RAiDs on national level. First involves e.g. the introduction and minting of RAiD identifiers for research projects involving numerous research entities and also disseminating RAiDs as national actor via APIs to other national/organizationl actors.

---

<sup>3</sup> Nikkanen, J. (2019). European Publication Information Infrastructure – metadata transfers in European context, <http://hdl.handle.net/11366/1226>.

<sup>4</sup> <https://github.com/EuroCRIS/CERIF-Core>

Enriching the research metadata on outputs, exploring workflows and graphs and implementing them in national CRIS-systems and providing national information to OpenAIRE, two components are crucially for the case study 1) Research Discovery Graph and PID Graph. The requirement gathering has started by discussing these with WP5 on how to actually utilize the PID Graph service and what the current situation relating to PIDs is at the moment in national CRIS systems. This is important not only for the metadata discoverability, but also also to interlinking different research entitites in a PID Graph. The latter is highly relevant for the case study in the sense of providing and interlinking as much of the research entities collected within the national CRIS systems.

## Conclusions

For both the EOSC RDGraph and OpenAIRE, the national CRIS systems form a valuable data source for information on the research entities themselves as well as the relationships between entities. Viceversa, national CRIS systems may be able to gain information on new connections between research objects from the EOSC RD- and PID Graphs. Moreover, this case study provides a framework for enhancing national level research assessment, monitoring, etc., while also enriching the EOSC RDGraph. The solution developed by this case study will serve as a template for integration of further national research information systems with the EOSC-Core.