Comments on "OPENPIIR Product Scope and Principles – recommendations of the ORG_ID Product Breakout Group"

About us

euroCRIS, founded in 2002, is an international not-for-profit association that brings together experts on research information in general and research information systems (CRIS) in particular.

The mission of euroCRIS is to promote collaboration within the research information community and advance interoperability through CERIF, the Common European Research Information Format. Our members include research performing organisations, funders, solutions providers from more than 40 countries throughout Europe and beyond. Areas of interest also cover the role of CRIS in the wider research eco system including virtual research environments, open access and open data repositories, as well as information access and exchange mechanisms, standards and guidelines and best practice for CRIS.

Our recommendation

OpenPIIR, as presented in the document, will be a useful element of global research information infrastructure. euroCRIS is willing to support this effort. This paper follows some suggestions given by Jan Dvorak, CERIF & CRIS Architectures Task Group leader and member of euroCRIS dated October 15th, 2017.

1. Aspects regarding Quality Assurance

We encourage the initiative to focus primarily on aspects of data quality assurance (QA) rather than technical aspects. Unlike in the case of persons respective PIDs, organisations are always "on the move." They might alter their names, addresses, size and structure, etc. The concept should explicitly state how to deal with these changes.

We have some concerns about a concept that delegates the responsibility for data quality mainly to the organisations which are represented in the database. There will be many organisations unaware that they even have an ID assigned to them. It is difficult to determine or define which organisational unit or even which person should be responsible for maintenance. How will the process be organised to ensure that only suitably qualified and responsible people interact on the quality issue?

We think that a solid model for a reasonable data quality procedure can be implemented by establishing national contact points. The international landscape of science is very diverse. With a certain degree of probability, it is possible to find people who have knowledge and experience of their own national scientific landscape but it is far more difficult to find organisations or people with this expertise at a global level.

The Org-ID-group has prepared a provider survey on current players with a focus on a global scope (such as ISNI, GRID) and commented on the shortcomings of these initiatives (Paper: "Organisation identifiers: current provider survey"). We would like to recommend extending the scope regarding national or regional initiatives. As we have come to learn from our members, many organisations have developed their own databases (e.g. for the purpose of funding, see for example the European Commission (EC), https://eacea.ec.europa.eu/erasmus-plus/actions/erasmus-charter_en), which represents an actual list (updated 27/10/2017) of about 5,000 European HEI, each with an agenda-specific identifier called "Erasmus Code"). These organisations also have a specific knowledge on technical, methodological and quality assurance aspects. Because this data is process-produced, it is usually of high (practical) quality.

EuroCris is a member organisation. Many of our members share knowledge on ID-topics and even on organisational Ids (as well as on their management and quality assurance). Assuming the project continues, we would be happy to compile a list of members who are able and willing to share their knowledge (or even data). For examples see "PTCRIS_OrgID - Portuguese Organisation Identifiers Authoritative System", https://doi.org/10.1016/j.procs.2017.03.046, or "Quality is the Product is Quality – Information Management as a Closed-loop Process (http://dspacecris.eurocris.org/handle/11366/332) that introduces the German OrgID-approach.

2. Aspects regarding technical questions

The data items about an organisation in the registry will probably include the following:

- the name of the organisation, possibly in multiple languages
- any other identifiers of the organisation, so that the ORG_ID can act as bridging identifier – for example ISNI (or Ringgold), Grid ID, Erasmus Code (see above)
- lifecycle dates such as the beginning and (for past organisations) the end of the organisation's existence
- provenance metadata for the above information incl. updates

Entries in the organisation registry will probably need to have links between each other. The semantics of such links can include:

- succession: acquisition, merger, split, replacement

- organisational hierarchy (when implemented)

The CERIF model covers all these aspects of information about organisations in the registry. The associated information exchange formats can serve as a good basis for the API of the openPIIR.

3. Other relevant initiatives

In the discussed document, it is highlighted that openPIIR is considered "not intentionally competitive". However, there is a product on the market that already sets standards today - the GRID initiative (https://www.grid.ac/).

For further planning, we recommend considering a model that uses the expertise of this initiative, but it is embedded in a new governance model that takes into account the important principles of the breakout group.

For determining the precise definitions of the data items of the registry, we recommend working with the CASRAI initiative, which has extensive experience in developing standard vocabularies in the broad context of research information and reporting.

Contact:

Juergen Guedler iD https://orcid.org/0000-0002-5026-0287, juergen.guedler@dfg.de

Standards Task Group Leader, euroCRIS

Head of Group Information Management

Deutsche Forschungsgemeinschaft (DFG) (German Research Foundation), Bonn, Germany