

# Cataloguing

## Introduction defining context and scope

The technological review for cataloguing covers a subset of the different concepts to be managed in catalogues, as seen in requirement section.

- **References catalogues:** persons and organisations, publications, research objects.
- **Federated catalogues:** datasets, resources, physical samples, procedures and software

Activity and event logs have not been considered in the technology review because the subject is not mature enough in RI and ICT to manage this information in catalogue yet. As a matter of priority, we focus on references and federated catalogues described above.

The review gives an overview of the software applications or systems and interface standards used for cataloguing related information.

## Change history and amendment procedure

The review of this topic will be organised by in consultation with the following volunteers: . They will partition the exploration and gathering of information and collaborate on the analysis and formulation of the initial report. Record details of the major steps in the change history table below. For further details of the complete procedure see item 4 on the [Getting Started](#) page.

Note: Do not record editorial / typographical changes. Only record significant changes of content.

Date	Name	Institution	Nature of the information added / changed

## Sources of information used

The standards considered are provided by the following bodies:

- W3C: <https://www.w3.org/>
- ISO: <http://www.iso.org/iso/home.html>
- OGC: <http://www.opengeospatial.org/>

RDA working groups (e.g. metadata) as well as domain specific standards (cerif, geoscienceML, ...) are also considered.

The tools are open-source or not. They might be desktop or server side (with web interface) software.

## Two-to-five year analysis

### References catalogues

- Persons and organizations: The most popular system for persons identification and cataloguing is currently OrcID. They are involved in THOR project which helps to connect together datasets, papers and researchers information. They are also working on organization cataloguing (<http://orcid.org/blog/2016/03/09/organisations-missing-link>).
- Publications management systems cited by RI are web of knowledge (<http://apps.webofknowledge.com>) and scopus (<http://www.scopus.com/>).
- For research objects no technology have been cited. Further investigation would be required.

### Federated catalogues:

- dataset catalogues are managed at RI level with CKAN in the Open Data world and RDA or geonetwork in the ISO and OGC. It sounds pragmatic and feasible to harvest existing CKAN and geonetwork in one CKAN central server (e.g. at EUDAT).
- resources, especially observation equipment are managed in dedicated systems at RI level, however two standards are popular to describe these items: SSN ontology from W3C and sensorML from OGC.
- physical samples are managed in dedicated systems. No common standard have been identified yet. However we can expect to type of solution from 1) biology discipline 2) from solid earth
- procedures and software: to be completed.

Overall solutions: CERIF proposed by EPOS provides an overall conceptual model for managing the above information.

To BE completed

## State of the art

## Trends

**Subsequent headings for each trend (if appropriate in this HL3 style)**

Problems to be overcome

**Sub-headings as appropriate in HL3 style (one per problem)**

Details underpinning above analysis

Sketch of a longer-term horizon

Relationships with requirements and use cases

Summary of analysis highlighting implications and issues

Bibliography and references to sources