

# A metadata repository (MDR) for clinical study objects

Wednesday, 8 May 2019 11:45 (15 minutes)

In order to keep track with evidence generation in medicine it is necessary not only to have access to results from clinical studies in publications but also to the individual participant data as well as all related study documents (e.g. study protocol, statistical analysis plan, case report form). Data and document sharing is more and more propagated, however, the researcher is faced with a bewildering mosaic of possible source locations and access modalities. There is an urgent need to develop a central resource that can catalogue all the diverse data and documents associated with a clinical study and make that information searchable by a central web portal. In the EU H2020-funded project eXtreme DataCloud (XDC; grant agreement 777367) such a portal is currently developed under coordination of ECRIN-ERIC (European Clinical Research Infrastructure).

## Methods

The use case description, covering goals, requirements, data sources, targeted metadata schema, data structures and user interaction and design, is provided in the XDC project description in confluence. The MDR portal will be integrated in the XDC infrastructure and will be based upon the following components:

- Importing metadata from existing registries and repositories
- Mapping of imported metadata to the ECRIN metadata standard
- Pumping of metadata via restful-API to OneData
- Providing functionality for discoverability studies and data objects by INFN (Elasticsearch)
- Developing GUI by OneData

## Results

As a first step the ECRIN metadata schema for clinical research based upon DataCite has been updated (1). So far metadata from 7 data sources have been imported (CT.gov, PubMed, WWARN, Edinburgh DataShare, BioLINCC, ZENODO, Data Dryad), covering more than 500000 records from clinical studies. The imported data sources are stored as JSON objects and in relational DB form on the test bed server at INFN, Bologna. The metadata acquired have been mapped to the ECRIN metadata schema using standard JSON templates. Currently under way are the upload of the mapped metadata into OneData and the provision of the search functionality by INFN.

## Conclusions

All preparatory work for the MDR has been performed, integration into the XDC infrastructure has been started. It is planned to have the first demonstrator with full functionality in April 2019, ready for testing. If successful, the MDR will be introduced as data resource into the European Open Science Cloud (EOSC).

## Type of abstract

Presentation

## References

1. Canham S, Ohmann C: A metadata schema for data objects in clinical research, *Trials* 2016; 17:557

**Primary author:** Mr GORYANIN, Sergei (ECRIN)

**Co-authors:** Dr ITALIANO, Alessandro (INFN); Prof. OHMANN, Christian (ECRIN); Dr DONVITO, Giacinto (INFN); Dr KUDZIA, Jakub (ACK Cyfronet AGH); Dr DUTKA, Lukasz (ACK Cyfronet AGH); Dr BATTAGLIA, Serena (ECRIN); Dr NICOTRI, Stefano (INFN); Dr CANHAM, Steve (ECRIN)

**Presenter:** Mr GORYANIN, Sergei (ECRIN)

**Session Classification:** Scientific and technical updates