GEDE DO Participation Template

## Name and Institution

Wolfgang Kuchinke, Heinrich-Heine University Duesseldorf, Germany

## Email and Telephone Number

wolfgang.kuchinke@uni-duesseldorf.de, Tel: (0049)2118116142

## Names of Collaborators

Christian Ohmann, Steve Canham

## Research Community

ECRIN

## Major Goals of the Ongoing Infrastructure work

The aim is to make clinical trials data and associated documents findable and sharable for further reuse by data re-analysis and meta-analysis. For this purpose the different DOs of a clinical trial, like study protocol, consent agreement, statistical analysis plan (SAP), publications, data sets, etc., are linked together, provided with metadata and identifiers and stored in suitable data repositories to make them findable and accessible. Because the many digital objects of a clinical trial are distributed throughout different repositories and are accessible or not under different conditions, a DO based infrastructure is needed to minimize search and access efforts and improve object discovery and re-use.

## Potential of the DO Concept for your Work

After the conclusion of a clinical trial, it is still too often the case that raw data is archived inaccessibly and only statistical summary results are published. In addition, the associated documents, like study protocol, data management plan, coding tables, statistical analysis plan, registration document, preliminary results, publications, are stored in different repositories and data bases. Thus they are partly accessible (publications) or often not accessible at all (study protocol, data management plan) and sometimes not even findable.

In this situation, the DO concept plays a major role in the infrastructure that is developed by ECRIN to support the linking and sharing of clinical trials objects. By creating metadata registries and clinical data repositories that are linked to existing repositories ECRIN is building the foundation of a data sharing infrastructure employing the DO architecture. As a first step, the multitude of clinical trials objects must be linked and amended by persistent identifiers and metadata records. A metadata schema was proposed that is based on the DataCite and contains extensions to cover the needs of clinical research [1]. This clinical research specific metadata schema will be integrated into the clinical trials data repository of ECRIN and be part of the developed Clinical Trial Metadata Registry. Links will be created to existing repositories that contain data and documents of clinical trials, like PubMed, ClinicalTrials.gov, Edinburgh DataShare, BioLincc, vivli, Yoda, Dataverse, CSDR, etc., and to general repositories, like Zenoto, B2SHARE and B2FIND [2].

In our concept, persistent identifiers are attached to all DO of clinical trials, including data sets, documents, software versions, algorithms, etc. Corresponding metadata allow to bring together all DOs of a single trial that are distributes in different repositories to them searchable. Metadata will include details about the sensitivity of data and access conditions of the data sources, to inform the searching person about access requirements and constraints. In this manner, a DO based approach to data sharing can open a way towards an open and controlled way of data provision and sharing; it can also build trust into the data sharing process.

**References:**

[1] Canham S, Ohmann C. A metadata schema for data objects in clinical research. Trials. 2016;17(1):557. doi:10.1186/s13063-016-1686-5

[2] Kuchinke W, Apweiler S. The use of the EUDAT repository to store clinical trials in a secure and compliant way. Available online: www.eudat.eu/communities/the-use-of-the-eudat-repository-to-store-clinical-trials-in-a-secure-and-compliant-way