



Contribution ID: 49

Type: Poster

Reproducible Clinical Research Necessitates Metadata Harmonization and Standardization

Metadata is structured information that describes data objects. It gives the data user the necessary context to extract information from the data. In research, metadata is essential for data quality checks, interpreting findings, and reproducing experiments. Currently, each institute or group has its own catalog of metadata they require for experiments and subsequent analyses. This makes data integration from different research groups almost impossible and hinders the replication of experiments. With the publication of the FAIR (findable, accessible, interoperable, and reusable) data principles, FAIR data as well as data and metadata standards receive increasing recognition. However, many researchers are unaware of the possibilities for structuring their metadata and how to enrich metadata from data providers or other third parties. In addition, clinical data pertaining to human subjects comes with its own ethico-legal challenges regarding privacy and security. Therefore, we want to highlight community standards for minimal clinical metadata harmonization and standardization that are applicable across a wide range of biomedical research disciplines.

Applying our experience from developing the German Human Genome-Phenome Archive metadata model, we collected community standards that serve to enhance data and metadata quality. These cover minimal reporting standards, various ontologies, and other best practice guidelines from clinical research and sequencing applications. As many scientists do not have to deal with legal challenges arising from human-related data, we additionally want to shed light on possible issues and offer workarounds that are GDPR-compliant and still enable fair data collection and sharing.

Standardization and harmonization of data are key steps during all steps of data collection and processing. Educating researchers about data and metadata standards in clinical research fields counteracts the impending reusability crisis and increases overall data quality.

Primary authors: SCHULTZE, Joachim L. (DZNE, Universität Bonn); MAUER, Karoline (DZNE); ULAS, Thomas (Universität Bonn, DZNE)

Presenter: MAUER, Karoline (DZNE)

Session Classification: Posters