Contribution ID: **13** Type: **not specified**

Marina Pozhidaeva (COMPBC): MeSyTo - Advancing Data Interoperability in Toxicology and Pharmacology Through Standardized Metadata and Ontologies Data Science and Modelling

Wednesday 7 May 2025 11:30 (15 minutes)

Author: Marina Pozhidaeva

Abstract: "Tackling chemical pollution as part of the Triple Planetary Crisis requires data science-based solutions that enable comprehensive environmental monitoring, risk assessment, and policy development. However, the lack of standardized and interoperable metadata in toxicology and pharmacology hinders the effective integration and reuse of critical data. The MeSyTo project aims to improve toxicological and pharmacological data interoperability and machine readability by developing standardised metadata frameworks. The project addresses inconsistencies in metadata annotations that hinder the integration and reuse of data in toxicology and pharmacology. By adopting the FAIR (Findable, Accessible, Interoperable, Reusable) principles, MeSyTo will create structured metadata to document exposure conditions, experimental settings and workflows for omics data from collection to storage in the data repositories.

Our key accomplishments include the development of a comprehensive metadata catalog through the integration and harmonization of metadata elements from the OECD Omics Reporting Framework, various omics data repositories, and wet lab protocols. Furthermore, we have established a domain-specific ontology by defining core classes and aligning them with external ontologies to ensure semantic consistency and interoperability across datasets.

Current efforts focus on validating the domain-specific ontology, developing constraints using SHACL for case-specific metadata annotation, and extending its application to multi-omics data integration. The overarching goal is to enhance reproducibility, data sharing and reuse, and to foster collaboration within Helmholtz and beyond, ultimately empowering researchers to tackle the grand challenges of chemical pollution and its impact on human and ecosystem health."

Session Classification: Session II: Data science and modelling , Chair: tbd