TOWARD ENHANCING INTEROPERABILITY OF TOXICOGENOMICS DATABASES

CONTACT DETAILS:
RDA IG: TOXICOGENOMICS INTEROPERABILITY IG
NAME: DIANA HENDRICKX
ORGANISATION: MAASTRICHT UNIVERSITY, THE NETHERLANDS
EMAILD.HENDRICKX[AT]MAASTRICHTUNIVERSITY.NL

Objectives:

The objective of this interest group is to prepare for a working group, who will identify a roadmap toward toxicogenomics database interoperability. Toxicogenomics is the application of genomic technology to study the impact of environmental chemicals on human health. The need for alternatives to animal testing has resulted in the expansion of toxicogenomics approaches, data types, and data collections. Much of these data are stored in databases such as diXa, Chemical Effects in Biological Systems (CEBS), DrugMatrix, Comparative Toxicogenomics Database (CTD), TG Gates and Connectivity Map. To guide the direction of a future roadmap towards database interoperability, user stories have been solicited from existing stakeholder groups. Analysis of the collected user stories will be used to produce a number of use cases and other requirements of the roadmap with the goal of meaningfully improving toxicogenomics data sharing, access and analysis.

On-going activities:

The Toxicogenomics Interoperability Interest Group (IG) was established following the June 26-27 "Workshop on Identifying Opportunities for Global Integration of Toxicogenomics Databases" at the National Institute for Environmental Health Sciences, North Carolina, USA. Researchers, data stewards, and managers from the USA and European Union presented their work, and shared visions of data needs for scientific advancement. Focused discussion described the scientific benefits of integrating datasets. To garner a better understanding of the need, a broad assessment of the current uses of toxicogenomics data within the research community was needed and would best be conducted within the structure of an RDA IG. The Toxicogenomics Interoperability IG is dedicated to serving toxicogenomics advancement by pursuing the RDA vision of making data work across barriers.

Foundational to the creation of a roadmap for toxicogenomics interoperability is an e-questionnaire deployed via outreach to database users and stakeholder groups in the field. Researchers were asked to describe their use of toxicogenomics data, the limitations of the data, and which improvements that would advance their own research mission. In this way user stories were collected, and will be analyzed to identify functional use cases to inform the creation of a data sharing and exchange roadmap.

Results:

While outreach is ongoing to capture the state of the field, e-questionnaire responses have described common data uses as model creation, reference data or data analysis. Perceived barriers as well as solutions to data use were heterogeneous, although the most commonly cited lacking data availability, poor metadata or data records, and study design differences. Overwhelmingly respondents of the survey indicated that the ability to combine relevant datasets would be the one most significant improvement to data use in the field. Additionally, there is a reported lack of availability of data and metadata. There is a need for better standardization and a more unified study design. Data are difficult to interpret. The quantity of the data is often insufficient. Data are not available in a central location and

there is lack of database interoperability. There is also uncertainty about the data quality. Better standardization, increase in publicly available data and a centralized repository are indicated as the most significant achievements that could be made when current barriers are overcome.

Next steps needed to better understand data use are to identify additional stakeholders to assure inclusion of varied perspectives, and to document the primary toxicogenomic data-streams to inform any future requirements.

URL: https://www.rd-alliance.org/internal-groups/toxicogenomics-interoperability-ig.html