

A TOXICOLOGICAL ONTOLOGY FOR THE OECD QSAR TOOLBOX

Romualdo Benigni¹, Olga Tcheremenskaia¹, George Fotakis² and Joop Deknecht³

¹Istituto Superiore di Sanità, Environment and Health Department, Viale Regina Elena 299, Rome 00161, Italy; ²Computational Assessment Unit, European Chemicals Agency, Annankatu 18, P.O. Box 400, FI-00121 Helsinki, Finland; ³Environment, Health and Safety Division, Organisation for Economic Co-operation and Development 2, rue André Pascal - 75775 Paris

We report on a new ontological resource aimed at standardizing and organizing the chemical toxicological databases in the OECD (Q)SAR Toolbox software, which is freely downloadable from www.qsartoolbox.org. The main objective of the Toolbox is to allow the user to use (Q)SAR methodologies to group chemicals into categories and to fill data gaps by read-across, trend analysis and (Q)SARs. In order to group chemicals successfully, robust methods are required which provide mechanistically relevant approaches to form categories. For mutagenicity, robust categories can be formed using different profilers relevant to genetic toxicity:

- DNA binding by OASIS;
- DNA binding by OECD;
- DNA alerts for AMES, MN and CA by OASIS;
- Carcinogenicity (genotox and nongenotox) alerts by ISS;
- in vitro* mutagenicity (Ames test) alerts by ISS;
- in vivo* mutagenicity (Micronucleus) alerts by ISS.

In the module “Endpoints” the user can retrieve experimental results from the resident databases (e.g. Bacterial mutagenicity ISSSTY, Micronucleus ISSMIC, Genotoxicity OASIS, Toxicity Japan MHLW).

For the successful toxicity prediction all profilers and datasets need to be standardized and annotated to the common terminology. The definition of ontology and of controlled vocabulary is a crucial requirement for the interoperability between toxicology resources.

The project started in 2012 with ontology development for Carcinogenicity, Repeated Dose Toxicity and Reproductive/Developmental Toxicity. At the beginning several related resources have been identified, including already existing ontologies freely available at the Bioportal ontology depository (e.g. NCI Thesaurus, Clinical Terms Version 3, Mouse pathology and Mouse adult gross anatomy). The OECD harmonised templates available as xml schema have been used as the basis for the development of the ontologies. The Web Ontology Language (OWL DL) supported by the Protégé OWL editor has been used for the ontology implementation.

At the end of the Project each entry of the experimental databases included in the Toolbox software will be associated with the ontology using the OWL hierarchy relationships and restriction rules. A final goal of the Project is the introduction of the ontology as the basis of data exchange and harmonization within the OECD QSAR Toolbox for better integration and standardization of experimental data.

ACKNOWLEDGEMENTS

This work is funded by ECHA/OECD Program: “Multiple framework contract with re-opening of competition for the provision of scientific support services – ONTOLOGY” (ECHA/2011/25), Specific Contract 1 (ECHA/2011/125) and Specific Contract 2 (ECHA/2012/261).