

# Collaborative Ontology Design and Implementation for the Open-Tox Predictive Toxicology Framework

Olga Tcheremenskaia(A), Romualdo Benigni (A), Ivelina Nikolova (B), Nina Jeliaskova (B), Sylvia Escher (C), Helvi Grimm (C), Vladimir Poroikov (D), Alexey Lagunin (D) and Barry Hardy\* (E)

(A) Istituto Superiore di Sanità, Environment and Health Department, Istituto Superiore di Sanita', Viale Regina Elena 299, Rome 00161, Italy; (B) Ideacon Ltd, A. Kanchev 4, Sofia 1000, Bulgaria; (C) Fraunhofer Institute for Toxicology & Experimental Medicine, Nikolai-Fuchs-Str. 1, 30625 Hannover, Germany; (D) Institute of Biomedical Chemistry of Russian Academy of Sciences, 119121 Moscow, Russia; (E) Douglas Connect, Baermeggenweg 14, 4314 Zeiningen, Switzerland

## 1 ABSTRACT.

The goal of OpenTox [1] is to develop an interoperable predictive toxicology framework with multiple web services providing users access to distributed toxicological resources including data, computer models, validation and reporting.

OpenTox applications are based on use cases that satisfy user needs in predictive toxicology. Two examples are available as prototype services [1]: ToxPredict predicts a toxicity endpoint using a chemical structure input provided by the user, ToxCreate creates a predictive toxicology model from a user-provided dataset.

The definition of ontology and controlled vocabulary is extremely important to the construction of the OpenTox data infrastructure. It contributes to the necessary standardization and rational organization of data, describing and linking toxicological endpoints information in a formal way, suitable for being processed automatically by the software, with minimal human interaction.

We initiated the development of toxicological endpoints ontology using the DL species of the Web Ontology Language (OWL DL) and the Protégé OWL editor. Several OWL ontology classes are needed to describe different toxicological studies. Each attribute in a toxicological dataset can be associated with an entry in the ontology. The main OWL classes are “ToxicityStudyType”, “TestSystem” (includes subclasses such as strains, species, sex, route of exposure, etc), “TestResult” (includes subclasses such as toxicity measure, test call, mode of action, etc), and “Organs”. Our ontology currently possesses 5 toxicity study types: carcinogenicity, mutagenicity, *in vivo* micronucleus, repeated dose toxicity and aquatic toxicity. One of the more challenging ontology classes is the detailed description of organs starting from organs systems up to histological components. The Organs ontology is important for repeated dose toxicity and carcinogenicity endpoints. The OpenTox Toxicological Endpoint ontology is available for community use and feedback through the OpenTox website [2]. The collaborative work on the project is currently curated using the Collaborative Protégé Editor, with different access privileges assigned to OpenTox members. The ontology is additionally open for community reviews and annotations. We have

also created OpenToxipedia [3] as a collaborative resource for the entry and editing of toxicology terms using a Semantic Media Wiki, and supporting the peer review and discussion of terms. This activity is linked but separated from the more formal ontology development.

Besides the ontology for toxicological endpoints, a formal OWL representation of modelling algorithms is available on the OpenTox website [4]. This ontology describes different algorithms used in predictive toxicology, e.g. calculation of chemical descriptors. The Blue Obelisk [5] dictionary on cheminformatics algorithms has been used in the initial algorithm ontology development. We plan to extend this ontology to a full description of every algorithm used in OpenTox, including references, parameters and default values that will facilitate the introduction of new algorithm classes.

Following the ontology principles and best practices established by the Open Biological Ontology (OBO) [6] community, we try to reuse existing community ontologies that overlap with the required toxicological domains, building the more specific concepts required for toxicity study description when necessary. We are planning to import several ontological classes from relevant ontology development efforts, such as Ontology for Biomedical Investigations OBI [7], NCI Thesaurus [8], and SNOMED Clinical Terms [9]. We will submit our ontology as soon as it is ready to Bioportal [10] for dissemination and feedback.

## REFERENCES

1. <http://www.opentox.org/toxicity-prediction>
2. <http://www.opentox.org/partners/wp3/Toxicological%20Endpoints%20Ontology>
3. <http://www.opentox.org/opentoxipedia>
4. <http://www.opentox.org/dev/apis/api-1.1/Algorithms/?searchterm=ontology>
5. <http://blueobelisk.sourceforge.net/wiki/AlgorithmDictionary>
6. <http://www.obofoundry.org/crit.shtml>
7. <http://purl.obolibrary.org/obo/obi>
8. <http://ncicb.nci.nih.gov/core/EVS>
9. <http://ihtsdo.org>
10. <http://biportal.bioontology.org/>