OpenTox Predictive Toxicology Framework: toxicological ontology and semantic media wiki-based OpenToxipedia

Bio-Ontologies 2011

July 15 -16, 2011, Vienna, Austria

Olga Tcheremenskaia (Istituto Superiore di Sanità, Rome, Italy)

RomualdoBenigni (A), IvelinaNikolova (B), Nina Jeliazkova (B), Sylvia E.Escher (C), Helvi Grimm (C), Thomas Baier (C), Vladimir Poroikov (D), Alexey Lagunin (D), MichaRautenberg (E) and Barry Hardy (F)

(A) IstitutoSuperiore di Sanità, Rome, Italy
(B) Idea-consult Ltd, Sofia, Bulgaria
(C) Fraunhofer Institute for Toxicology &Experimental Medicine, Hannover, Germany
(D) Institute of Biomedical Chemistry of RussianAcademy of Sciences, Moscow, Russia
(E) In silico Toxicology, Basel, Switzerland
(F) Douglas Connect, Zeiningen, Switzerland

OpenTox - An Open Source Predictive Toxicology Framework

- OpenTox www.opentox.org, is funded under the EU Seventh Framework
- Implemented as open source REST web services
- Exchange of data and knowledge with ontologies (RDF, OWL-DL)
- Semantic interoperable computing architecture for the field of predictive toxicology.

In Silico Toxicology, Switzerland

Ideaconsult, Bulgaria

Istituto Superiore di Sanità, Italy

Technical University of Munich, Germany

David Gallagher, UK

Douglas Connect, Switzerland

Institute of Biomedical Chemistry of the Russian Academy of Medical Sciences, Russia

Albert Ludwigs University Freiburg, Germany

National Technical University of Athens, Greece

Fraunhofer Institute for Toxicology & Experimental Medicine, Germany

Seascape Learning & JNU, India

OpenTox

REACH: need for high quality toxicity data and predictive toxicology application

“Registration, Evaluation, Authorisation & Restriction of Chemicals”

European Union Legislation, 2007

Registration submitted by 22 Nov. 2010: 19,237

Chemicals pre-registered by 1 Dec. 2008: **143,000**

54M – 9M additional test animals

“...promote alternative methods for assessing hazards of substances.....animal testing as a last resort...” Geert Dancet, ECHA

Information on intrinsic properties of substances may be generated by using sources of information other than *in vivo testing:*
(Q)SARs ((Quantitative) Structure Activity Relationships), in vitro tests grouping of substances / category approach and read across approach.

The definition of ontology, and of controlled vocabulary is a crucial requirement in order to standardize and organize the chemical toxicological databases on which the predictive toxicology methods build on.

OpenTox

Toxicological data: need for standards

- Needs for data standards for automatic data integration
  - Example: **Carcinogenic Activity**

CPDBAS: Carcinogenic Potency Database
http://www.epa.gov/ncct/dsstox/sdf_cpdbas.html#SDFFields

ISSCAN: Chemical Carcinogens Database
http://www.iss.it/ampp/dati/cont.php?id=233&lang=1&tipo=7

Why we need a toxicological ontology?
-more efficient data– reducing the cost, maintenance and risk of the projectThe
-distinction among different synonyms, homonyms and related terms facilitates data export and import
-improvement of storage, exchange and use of information: unnecessary repetition of experiments can be avoided
-automatic reasoning
-avoiding errors
-intoducing restrictions (rules), etc

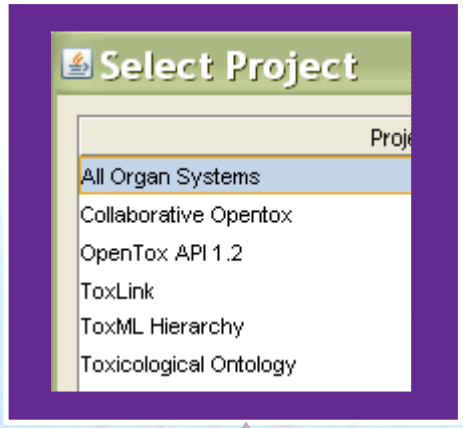
OpenTox

Collaborative Ontology Development: Collaborative Protege Server

OpenTox API 1.2 ontology

OpenTox Toxicological Data Ontology
5 toxicological endpoints

OpenTox Organs and Effects Ontology

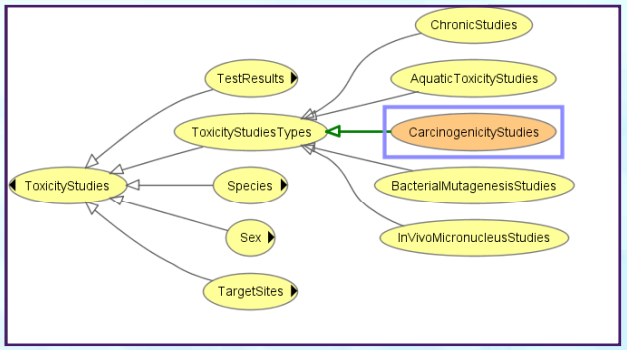


Collaborative OpenTox features, agorithms ontology

ToxML Hierarchy

ToxLink *in vitro* assays ontology (OpenTox-ToxCast) work in progress

OpenTox Toxicological Endpoint Ontology



•Why we need an ontology?

- Distributed services need to be able to "talk to each other", i.e. have a common understanding of endpoints, any type of property, methods, etc

•Methodology

- Starting from 5 toxicological endpoints
- Following the OBO Foundry principles
- Using the OWL-DL specie of the Web Ontolgoly Language

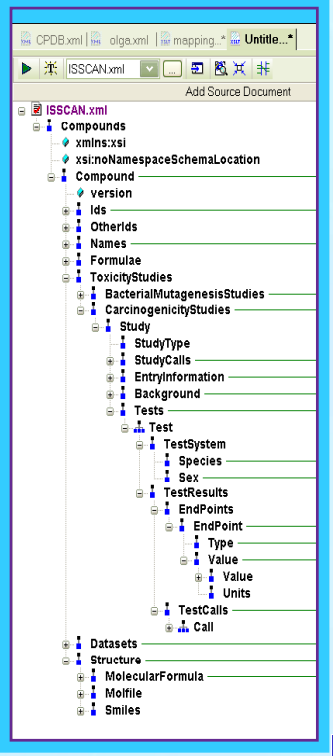
Other freely available resources: DSSTOX, INHAND, GoReni (ITEM), etc

Collaborative Protégé OpenTox Server

OpenTox Toxicological Data Ontology

Re-use of pieces or terms defined in neighboring ontologies (OWL and OBO)

ToxML scheme



Ontology For Toxicological Endpoints: overlapping with other biomedical ontologies

Chemical entities of biological interest (ChEBI)

Ontology for biomedical investigations OBI

NCI Thesaurus

Foundational Model of Anatomy (FMA)

Mouse Adult Gross Anatomy (MA)

Mouse gross anatomy and development (EMAP)

Units of measurement UO Ontology

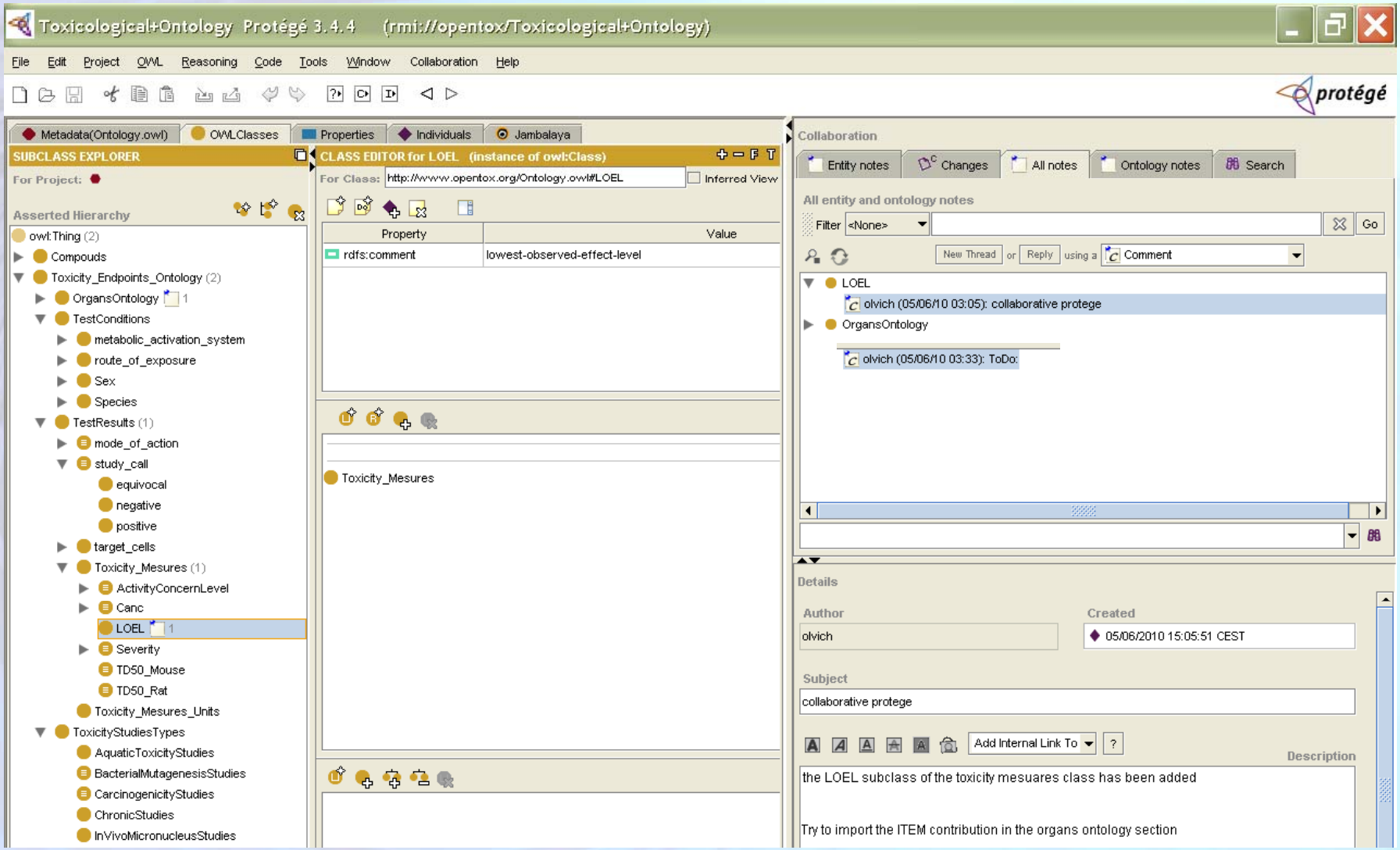
Chemical Information Ontology

CARO – Common Anatomy Reference Ontology

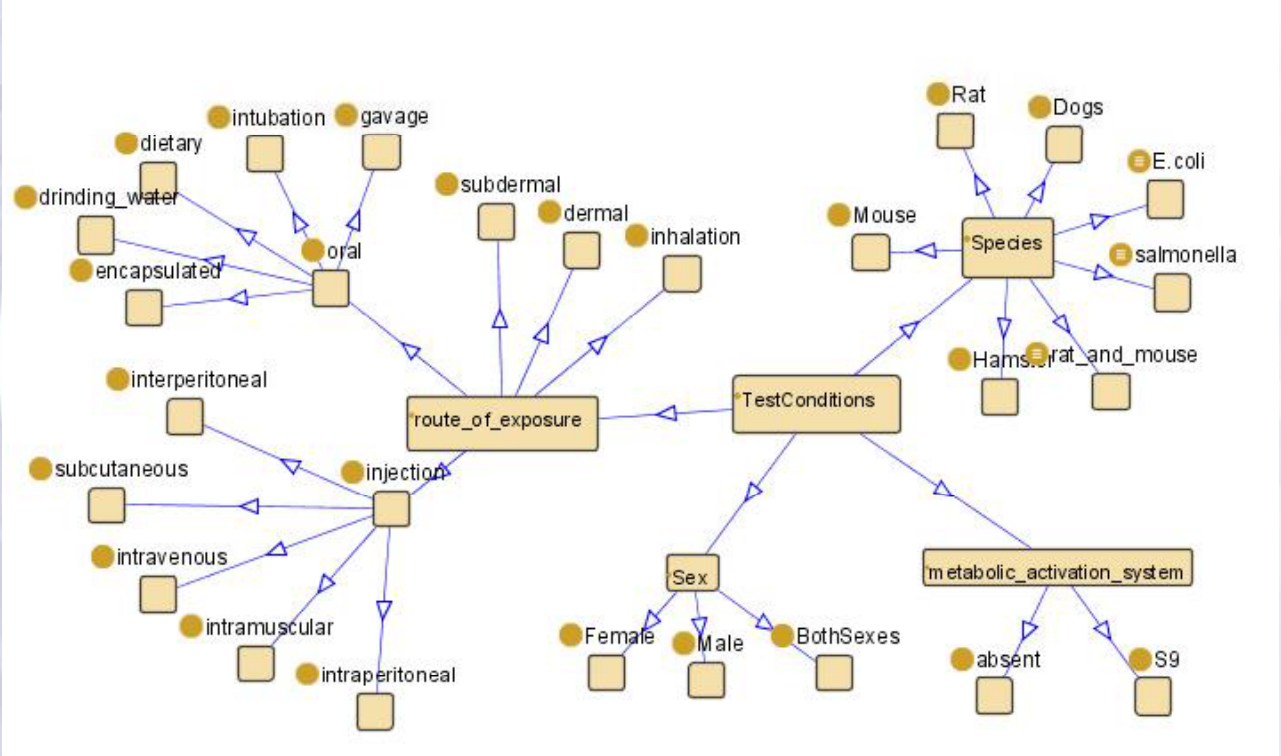
Ontology of scientific experiments (EXPO).

No systematic ontology for toxicological effects and predictive toxicology exists on the OboFoundry or the Bioportal Web sites

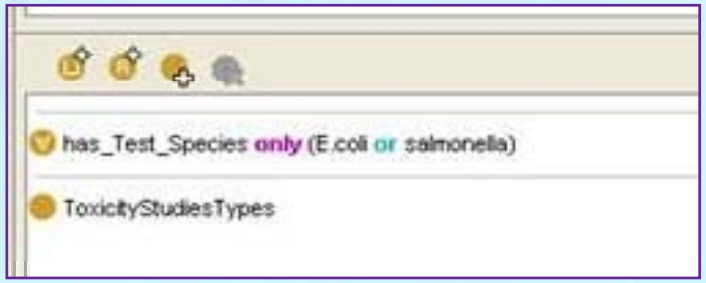
Collaborative Protègè: ToxicologicalOntology.owl screenshot



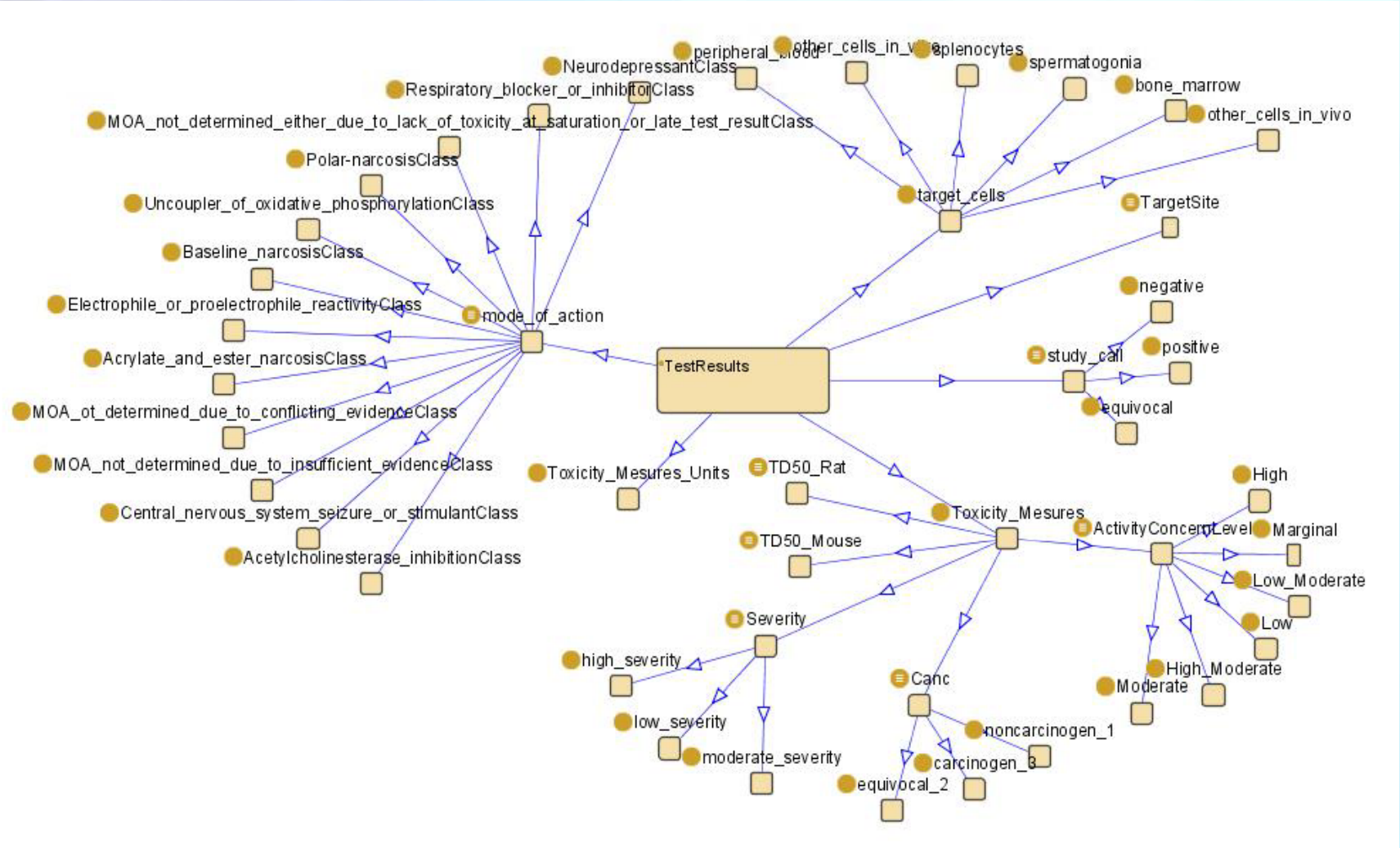
Main OWL Classes: TestConditions and TestResults



- Introduction of relationship between different classes, restrictions: e.g., introducing the property “has_Test_Species” limites the test species suitable for certain toxicity study.



Main Classes: TestConditions and TestResults



ToxML: conversion in OWL ontology

- **ToxML** (http://www.leadscope.com/toxml.php)
  - is a public initiative led by scientists at LeadScope, Inc
  - controlled vocabularies and XML scheme for storing chemical toxicity data

➢ Initial work to check out how costly would be to convert ToxML to an OWL ontology

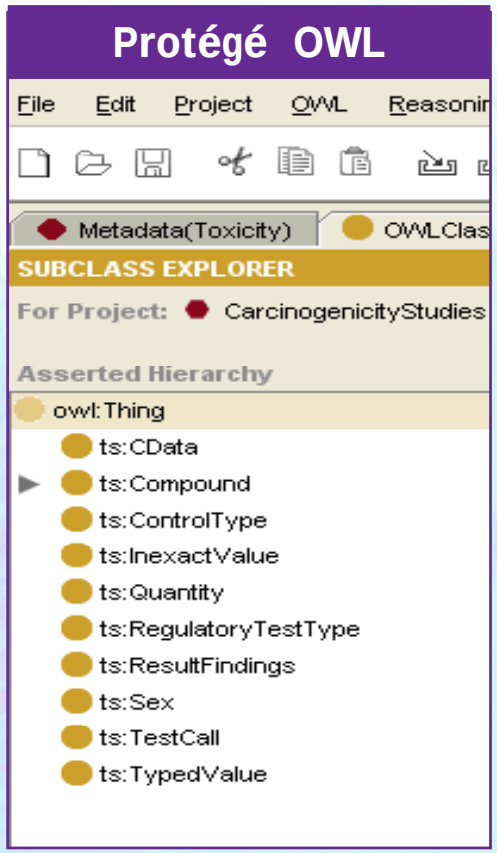
➢ Why? -to integrate the Leadscope databases in the OpenTox service
➢-to combine with data coming from different databases (e.g. for complex queries)

ToxML xsd schema

each element which doesn't have a type in the schema is converted to an OWL class

each element which has a type in the summary file is considered as a property

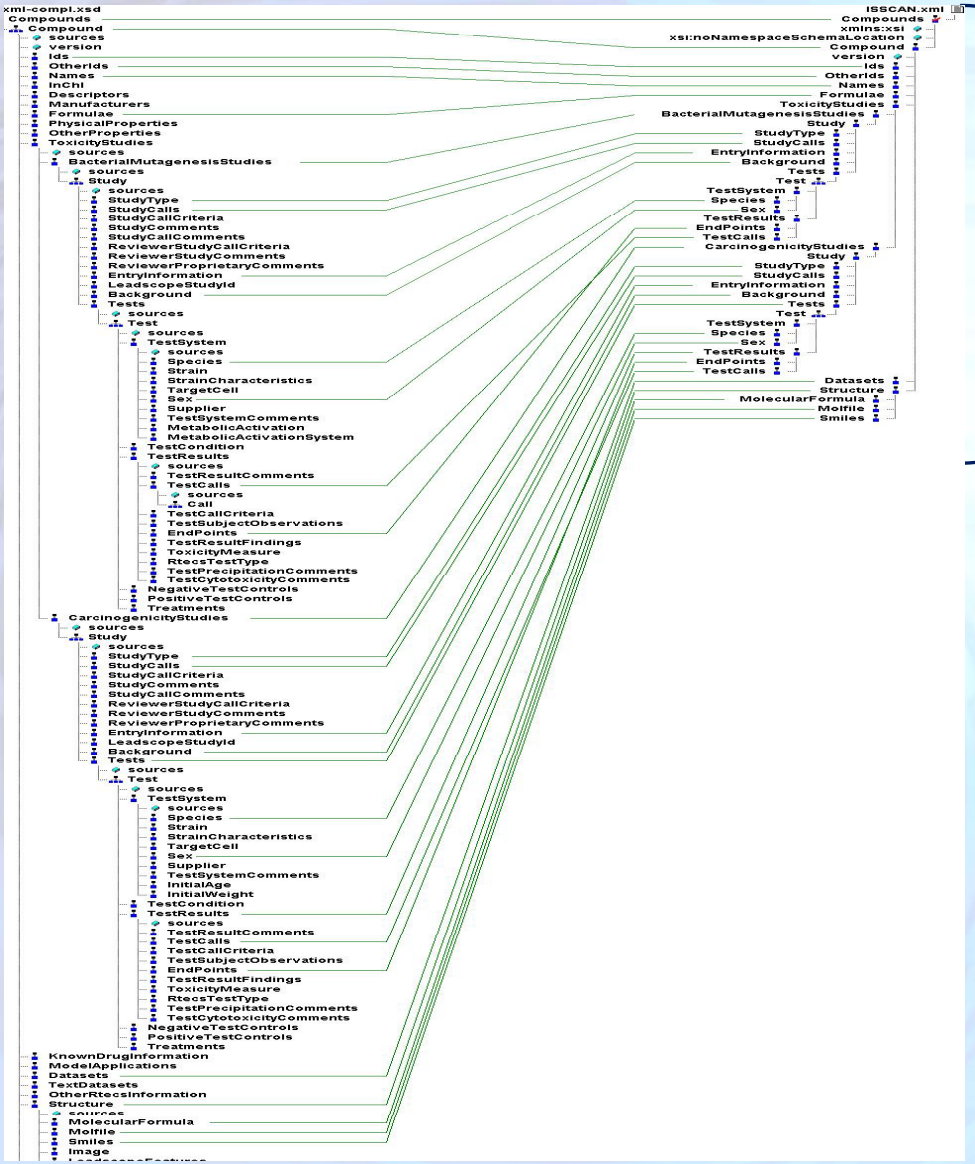
the parent relationships are kept



Some parts of the taxonomy may need to be reorganized

Needs for extensions: e.g. target sites, target cells, species are free text fields. Solutions: development extension in OWL, e.g. Organs Ontology has been developed importing of parts of the neighboring ontologies

Mapping of the ISSCAN entry - ToxML xsd scheme



CAS	67-66-3
Substance ID	4
ChemName	Chloroform
Synonyms	Formyl trichloride; methane trichloride; methenyl trichloride; Methyl trichloride; R 20; r 20 (refrigerant); Refrigerant R20; trichloroform; Trichloromethane
SAL	negative
Canc	positive
Reference	CPDB
TD50_Rat	262
TD50_Mouse	90.3
Rat_Male_Canc	positive
Rat_Male_NTP	ND
Rat_Female_Canc	positive
Rat_Female_NTP	ND
Mouse_Male_Canc	positive
Mouse_Male_NTP	ND
Mouse_Female_Canc	positive
Mouse_Female_NTP	ND
MolWeight	119.38
Formula	CHCl3
SDF file	Connection table
SMILES	ClC(Cl)Cl

OpenTox

Organs and Effects Ontology

- Contribution of **Fraunhofer Institute for Toxicology and Experimental Medicine (ITEM)**

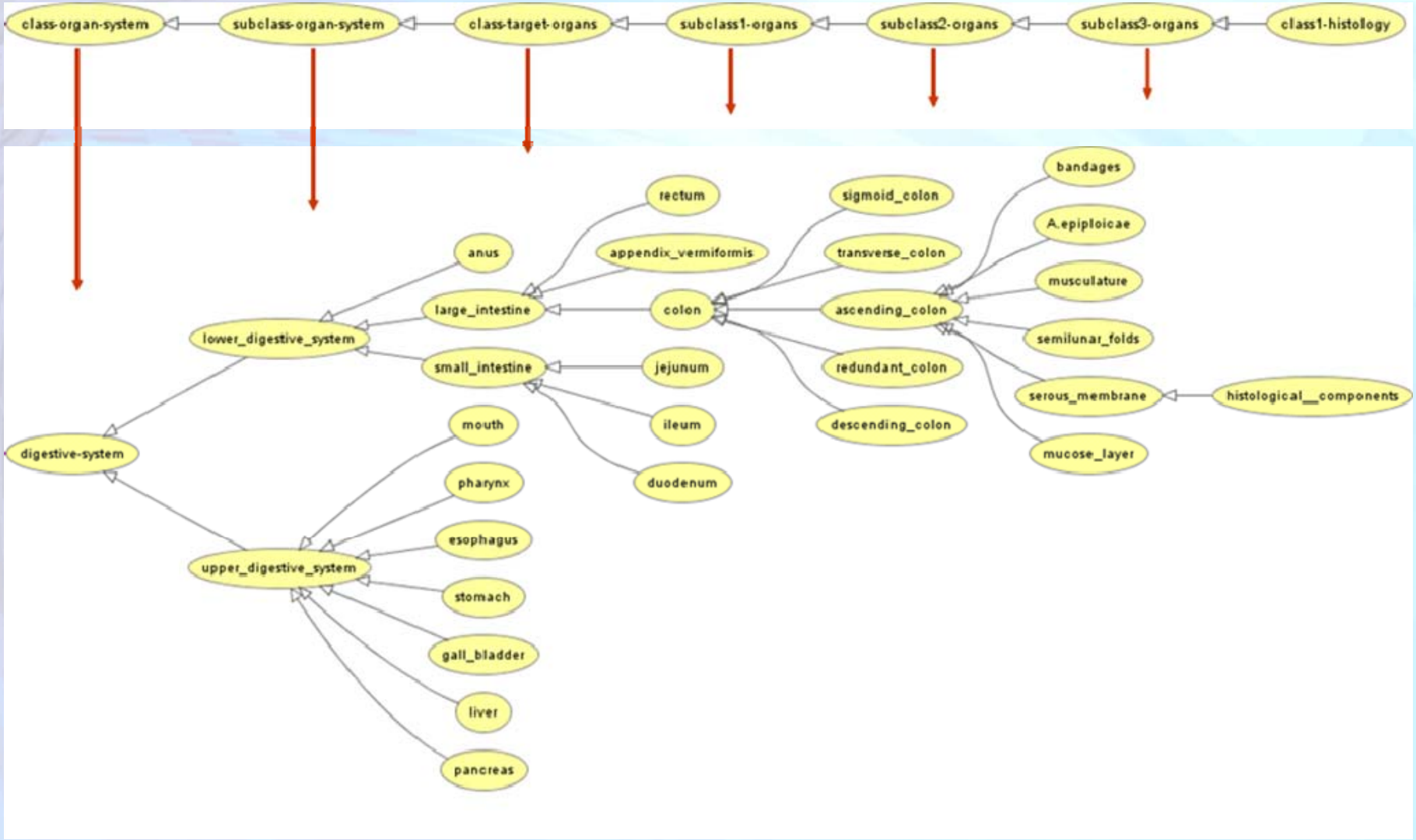
- Ontology Development on the base of **INHAND** (International Harmonization of Nomenclature ad Diagnostic Criteria for Lesions in Rates and Mice)

  Organs system (class) –Subclass organs system
  Target organs (class) - Targets organs (subclass 1 to N)
  Histopathology (class) -Histopathology (subclasses if needed)

- Key Features of the ITEM Organs Ontology:
  - more synonyms regarding organ systems, target organs and their subclasses
  - organs are more detailed, up to histological components
  - organs linked to different organ systems (e.g. ovaries)
  - reviewed by pathologists, who have been involved in the INHAND process
  - 12 organ systems fully described

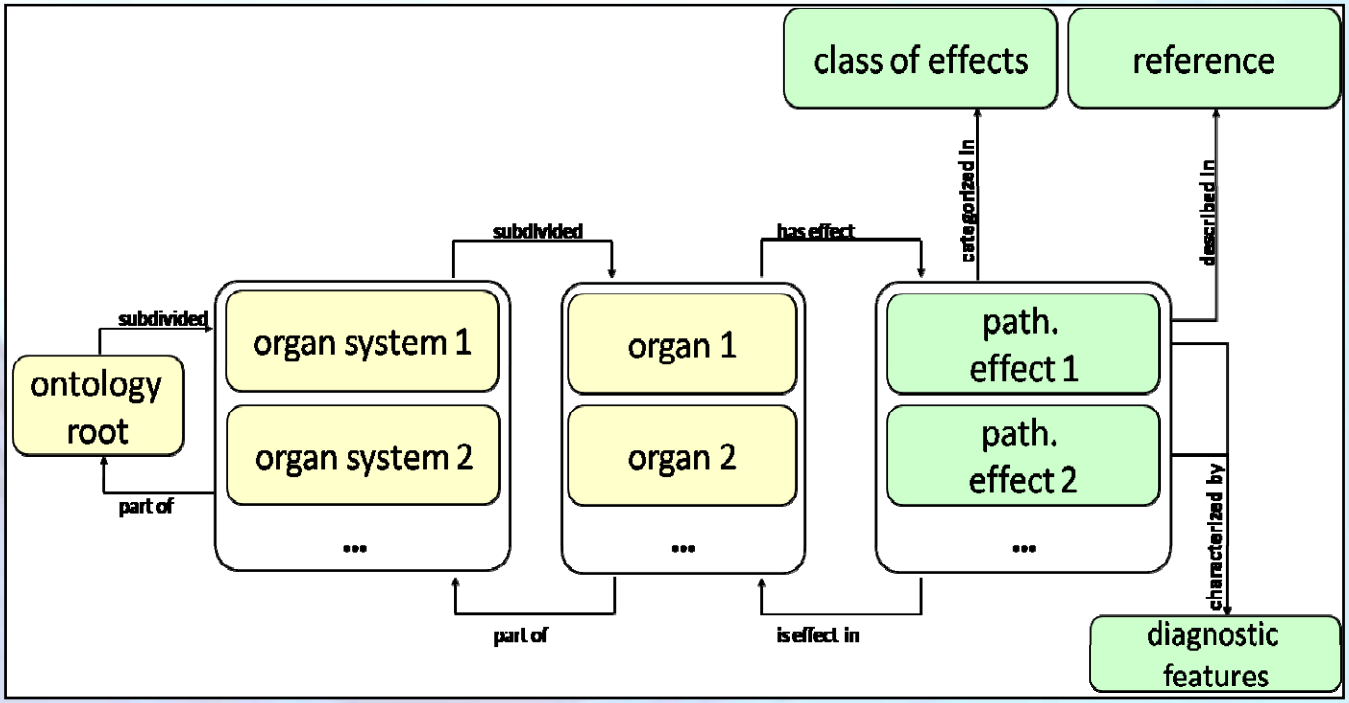
- Perspectives:
  - linking of the organ systems and their components with pathologic effects

Ontology For Target Organs: Digestive system example

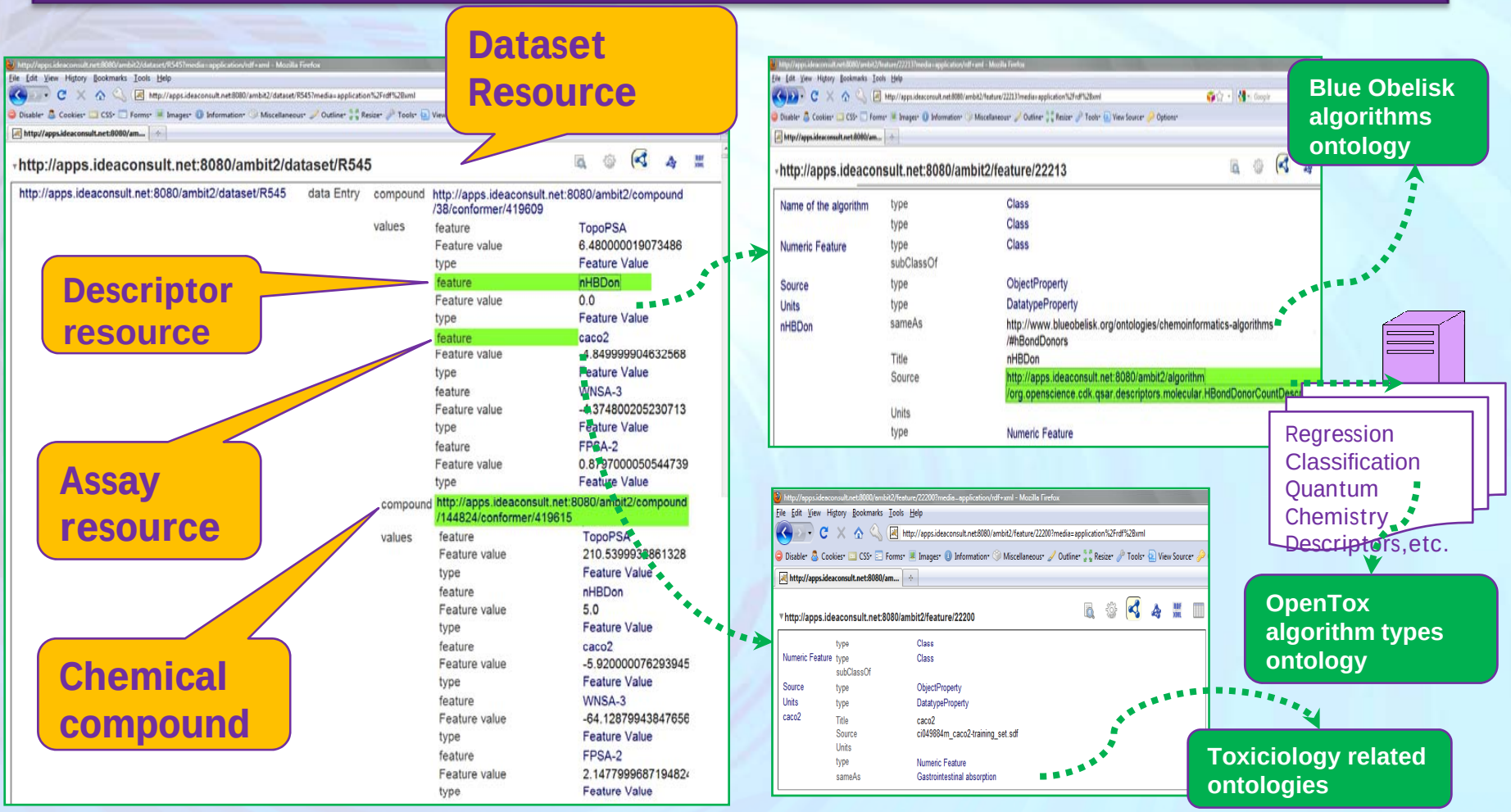


Structure of the combined Organs and Effects ontologies

The effects ontology consists of three main parts: classes of effects, linked to pathological effects, which are furtherlinked to detailed diagnostic features as agreed in the INHAND initiative. Its functionality has been initially developed for the respiratory tract.

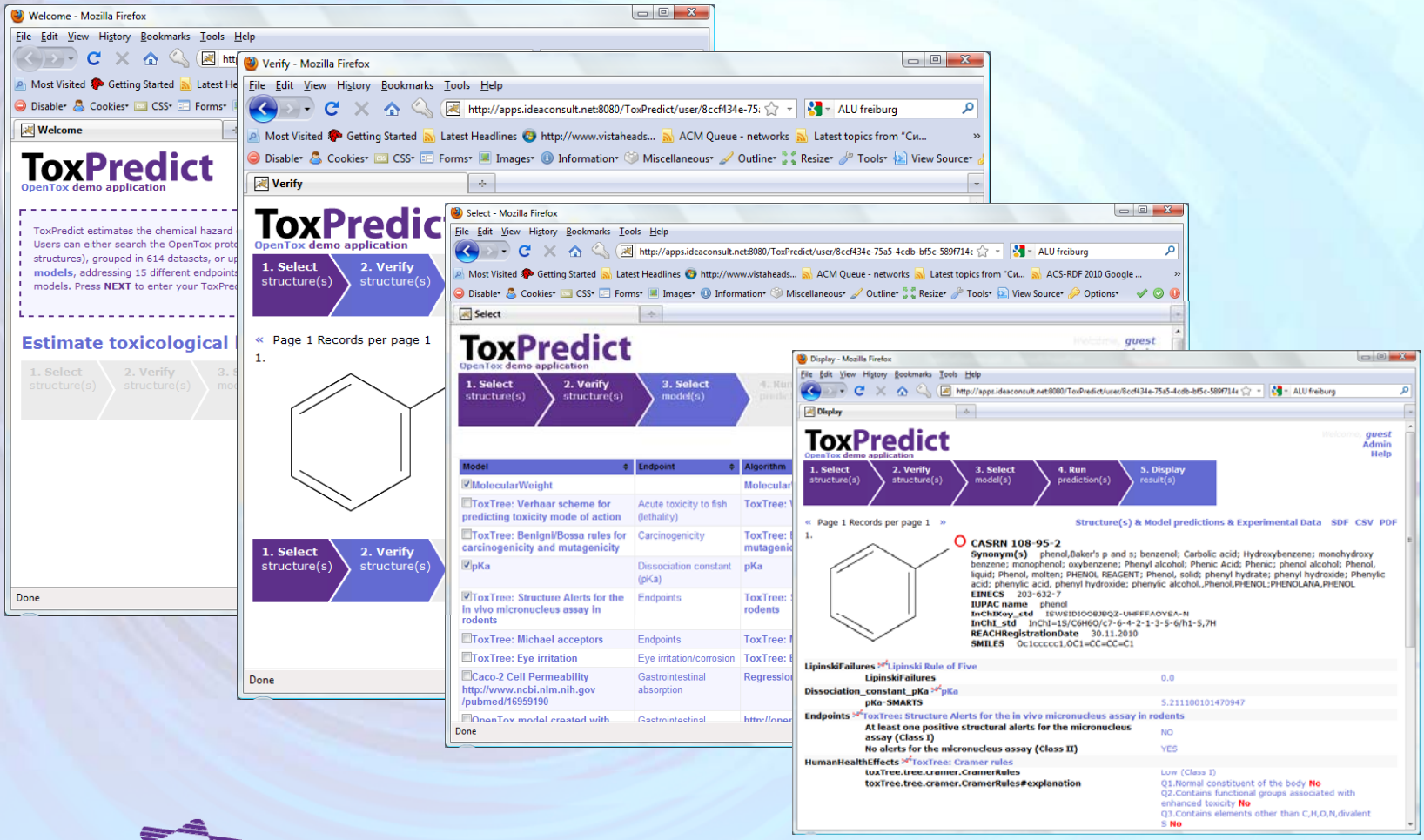


Linked resources and Ontology use: Compound, Algorithm, Model, Dataset, Features

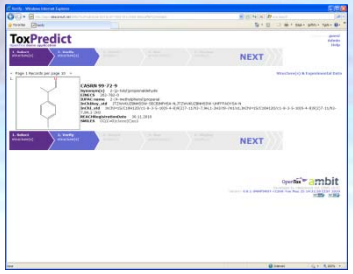


ToxPredict http://toxpredict.org

ToxPredict uses existing OpenTox models to estimate chemical compound properties. Developed and hosted by IdeaConsult partner.



ToxPredict and Ontology Service (behind the scenes)



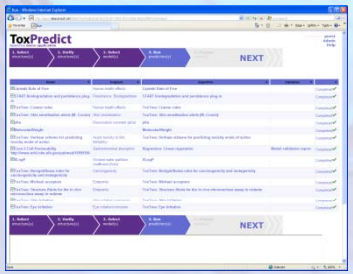
What prediction models are available? Is there a model for endpoint X?

ToxPredict Web Application

HTTP GET SPARQL query

OT Ontology Service

application/sparql-results+xml

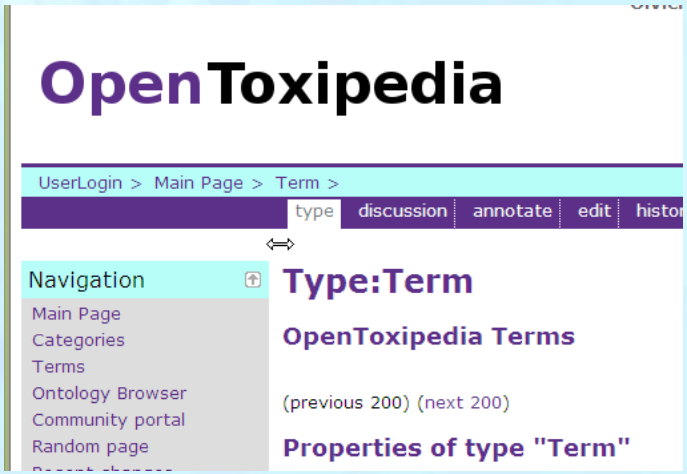


Here is the list of model URIs and related endpoints and algorithms in SPARQL format.

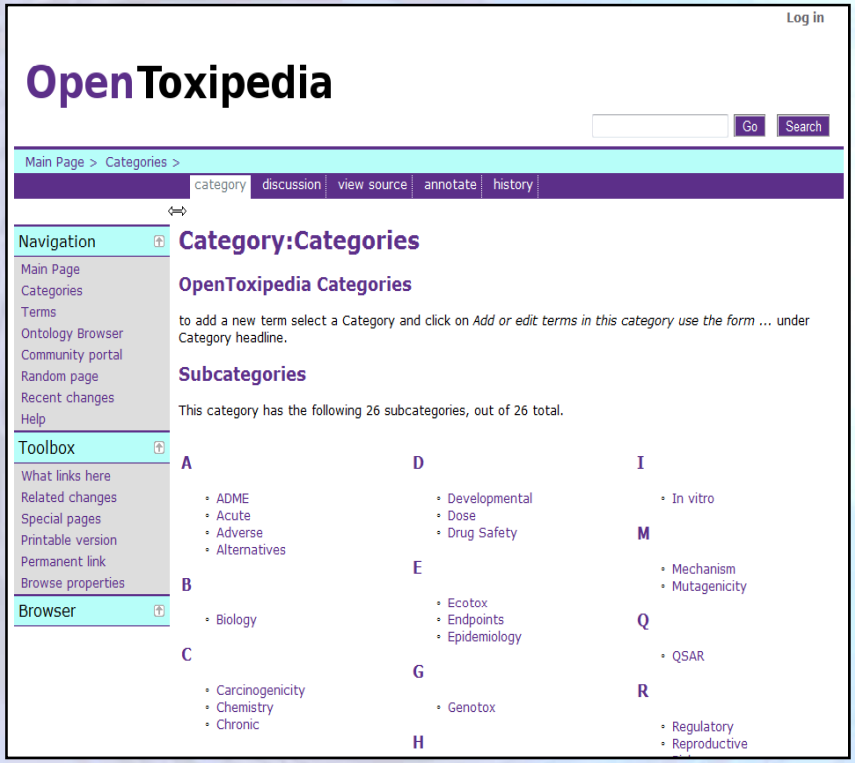
OpenToxipedia: community-based, predictive toxicology knowledge resource

- The OpenToxipedia organized by means of Semantic Media Wiki.
- http://www.opentoxipedia.org
  - Community based collaborative database
  - Automatically-generated lists.
  - Improved data structure, semantic terms anntotation
  - External reuse: RDF export and import in Protégé Ontology Editor

- **Creating, adding, editing and keeping terms used in toxicology terminology**
- **862 toxicological terms with description and literature references classified into 26 categories**
- **Transparency and scientific basis of information**
- **Curation by the OpenTox toxicologists**



OpenToxipedia: community-based, predictive toxicology knowledge resource



The terms can be browsed either by category or in alphabetical order.

Specialists in different toxicology fields are invited to take part in the creation and curation of OpenToxipedia.

The following rules for term management:
Add terms – any registered user (curators receive a message and decide what additions are approved and will become publicly available)
Edit description of terms – curators
Add remarks - any registered user (curators receive an alertmessage).

The particular importance of OpenToxipedia relies on the description of all the terms used in OT applications such as ToxPredict and ToxCreate.

Plan for Future Work

- Continue collaborative ontology development using the Collaborative Protégé server, covering more toxicological endpoints
- External collaboration with other related ontology projects
- Continue development of the ToxLink sub-project for *in vitro* assays ontology (in collaboration with the Computational Toxicology Research Program ToxCast that includes over 650 state-of-the-art rapid *in vitro* tests)
- Finish the conversion of the ToxML schema in the OWL
- Upload OpenTox Toxicological Endpoints Ontology linked to the Organs and Effects Ontology to the BioPortal Web Site
- OpenToxipedia: community based development

OpenTox Ontology Working Group

- **External collaboration coordinator**
  - Barry Hardy
- **Algorithms and features ontology**
  - Nina Jeliazkova,
  - Vedrin Jeliazkov,
  - Ivelina Nikolova
  - Christoph Helma
  - Tobias Girschick
  - Andreas Karwath
  - Georgia Melagraki
  - Sunil Chawla
  - David Gallagher
- **Collaborative Protégé administrator**
  - Micha Rautenberg
- **Toxicological Endpoint Ontology**
  - Olga Tcheremenskaia
  - (coordinator of the OpenTox Ontology Working Group)
  - Romualdo Benigni
- **Organs and Effects Ontology**
  - Sylvia Escher
  - Helvi Grimm
  - Thomas Baier
- **OpenToxipedia**
  - Alexey Lagunin
  - Sergey Novikov
  - Natalya Skvortsova

Thank you very much for your attention!