

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)



Romualdo Benigni (A), Ivelina Nikolova (B), Nina Jeliaskova (B), Sylvia E. Escher (C), Helvi Grimm (C), Thomas Baier (C), Vladimir Poroikov (D), Alexey Lagunin (D), Micha Rautenberg (E) and Barry Hardy (F)

(A) Istituto Superiore 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 Russian Academy of Sciences, Moscow, Russia

(E) In silico Toxicology, Basel, Switzerland

(F) Douglas Connect, Zeiningen, Switzerland



Bio-Ontologies 2011 SIG, ISMB/ECCB, July 15 -16, 2011, Vienna, Austria



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

Seascope Learning &
JNU, India

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.

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

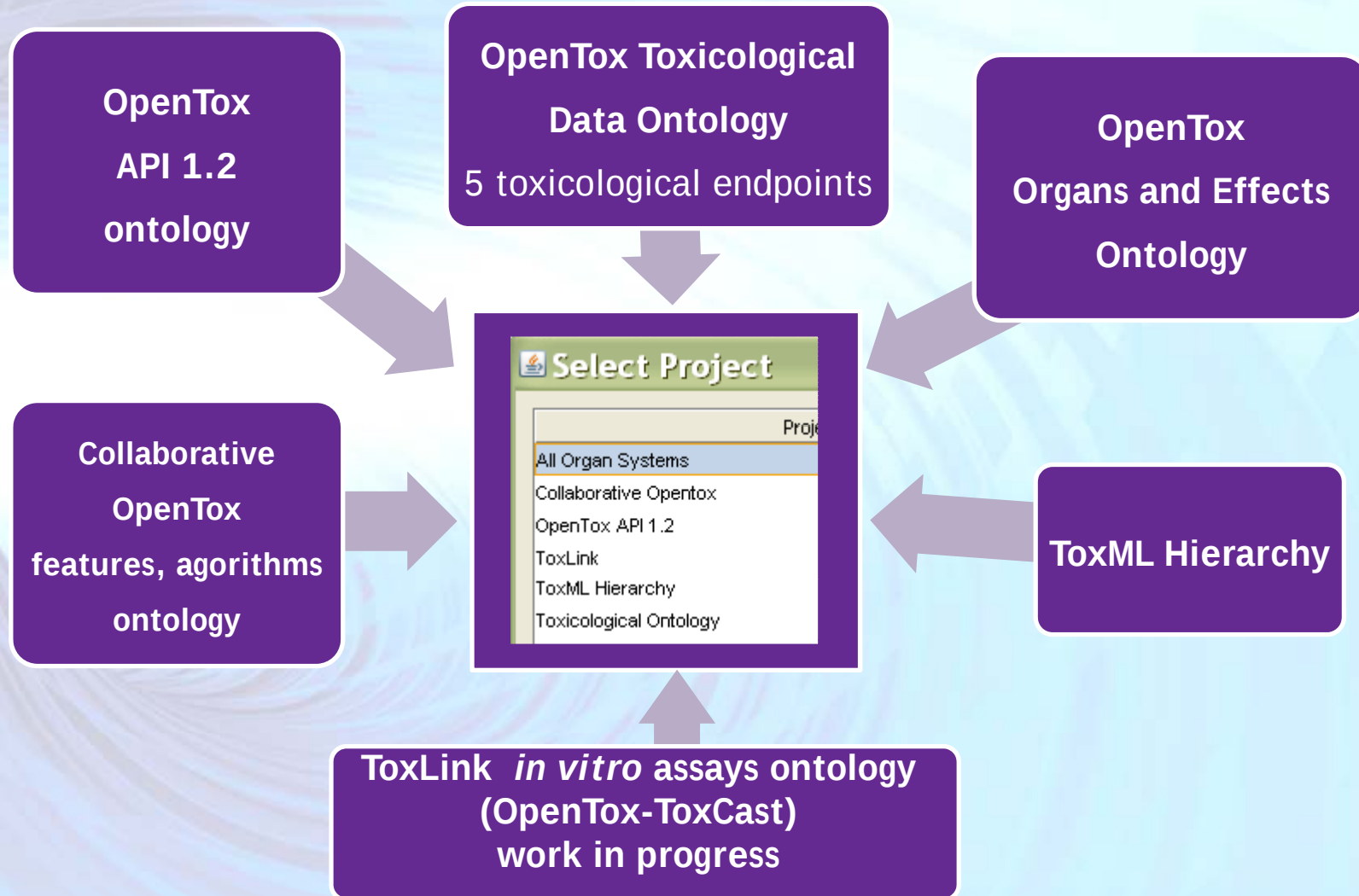
ISSCAN: Chemical Carcinogens Database
<http://www.iss.it/ampp/dati/cont.php?id=233&lang=1&tipo=7>



Why we need a toxicological ontology?

- more efficient data- reducing the cost, maintenance and risk of the project
- The
- 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
- introducing restrictions (rules), etc

Collaborative Ontology Development: Collaborative Protege Server



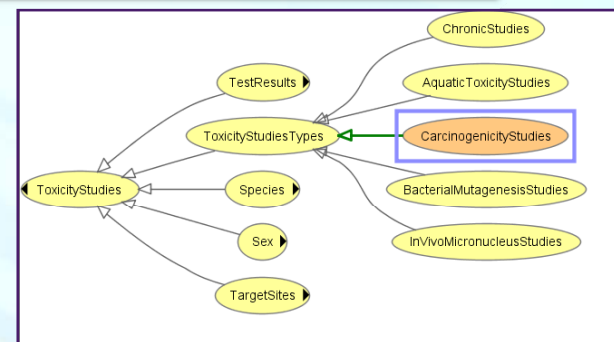
OpenTox Toxicological Endpoint Ontology

•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 Ontology Language



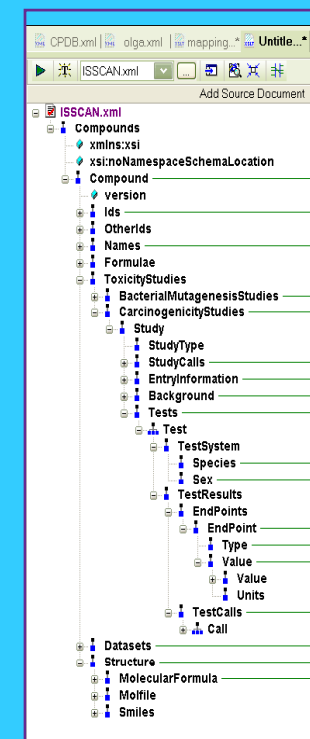
Other freely available resources: DSSTOX, INHAND, GoReni (ITEM), etc

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).

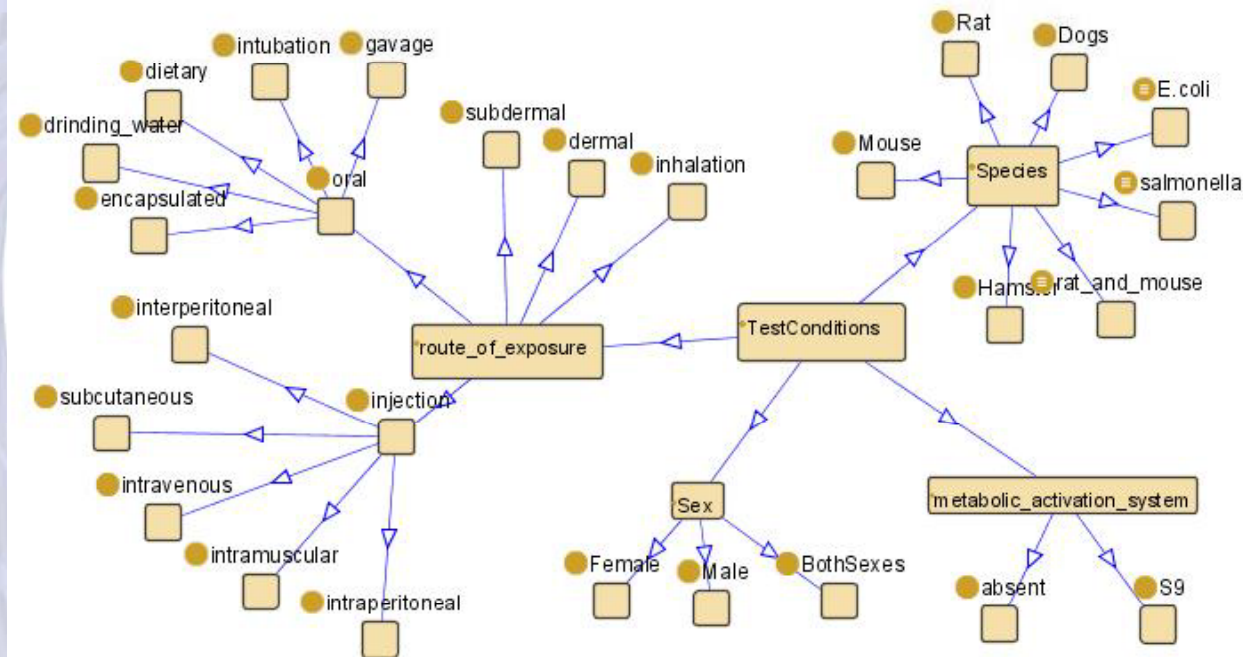
No systematic ontology for toxicological effects and predictive toxicology exists on the OboFoundry or the Bioportal Web sites

Collaborative Protégé: ToxicologicalOntology.owl screenshot

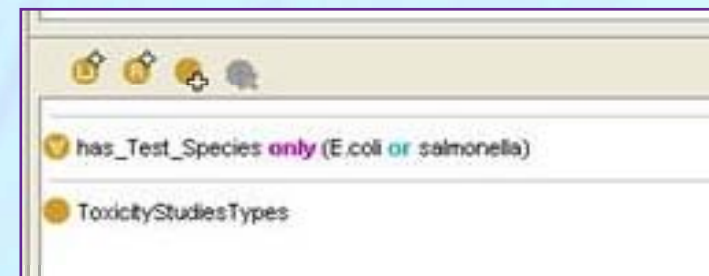
The screenshot displays the Protégé 3.4.4 application window titled "Toxicological+Ontology Protégé 3.4.4 (rmi://opentox/Toxicological+Ontology)". The interface is divided into several panels:

- Left Panel (SUBCLASS EXPLORER):** Shows the "Asserted Hierarchy" for the project. The hierarchy includes classes like `owl:Thing`, `Compounds`, `Toxicity_Endpoints_Ontology`, `OrgansOntology`, `TestConditions`, `metabolic_activation_system`, `route_of_exposure`, `Sex`, `Species`, `TestResults`, `mode_of_action`, `study_call`, `equivocal`, `negative`, `positive`, `target_cells`, `Toxicity_Mesures`, `ActivityConcernLevel`, `Canc`, `LOEL` (selected), `Severity`, `TD50_Mouse`, `TD50_Rat`, `Toxicity_Mesures_Units`, `ToxicityStudiesTypes`, `AquaticToxicityStudies`, `BacterialMutagenesisStudies`, `CarcinogenicityStudies`, `ChronicStudies`, and `InVivoMicronucleusStudies`.
- Central Panel (CLASS EDITOR for LOEL):** Shows the class editor for `LOEL` (instance of `owl:Class`). The "For Class" field is set to `http://www.opentox.org/Ontology.owl#LOEL`. The "Property" table lists `rdfs:comment` with the value `lowest-observed-effect-level`. Below the table, the `Toxicity_Mesures` class is visible.
- Right Panel (Collaboration):** Shows the "Collaboration" section with tabs for "Entity notes", "Changes", "All notes", "Ontology notes", and "Search". The "All entity and ontology notes" section displays a list of notes, including "olvich (05/06/10 03:05): collaborative protege" and "olvich (05/06/10 03:33): ToDo:". Below this, the "Details" section shows the "Author" as "olvich", the "Created" date as "05/06/2010 15:05:51 CEST", the "Subject" as "collaborative protege", and the "Description" as "the LOEL subclass of the toxicity mesuares class has been added" and "Try to import the ITEM contribution in the organs ontology section".

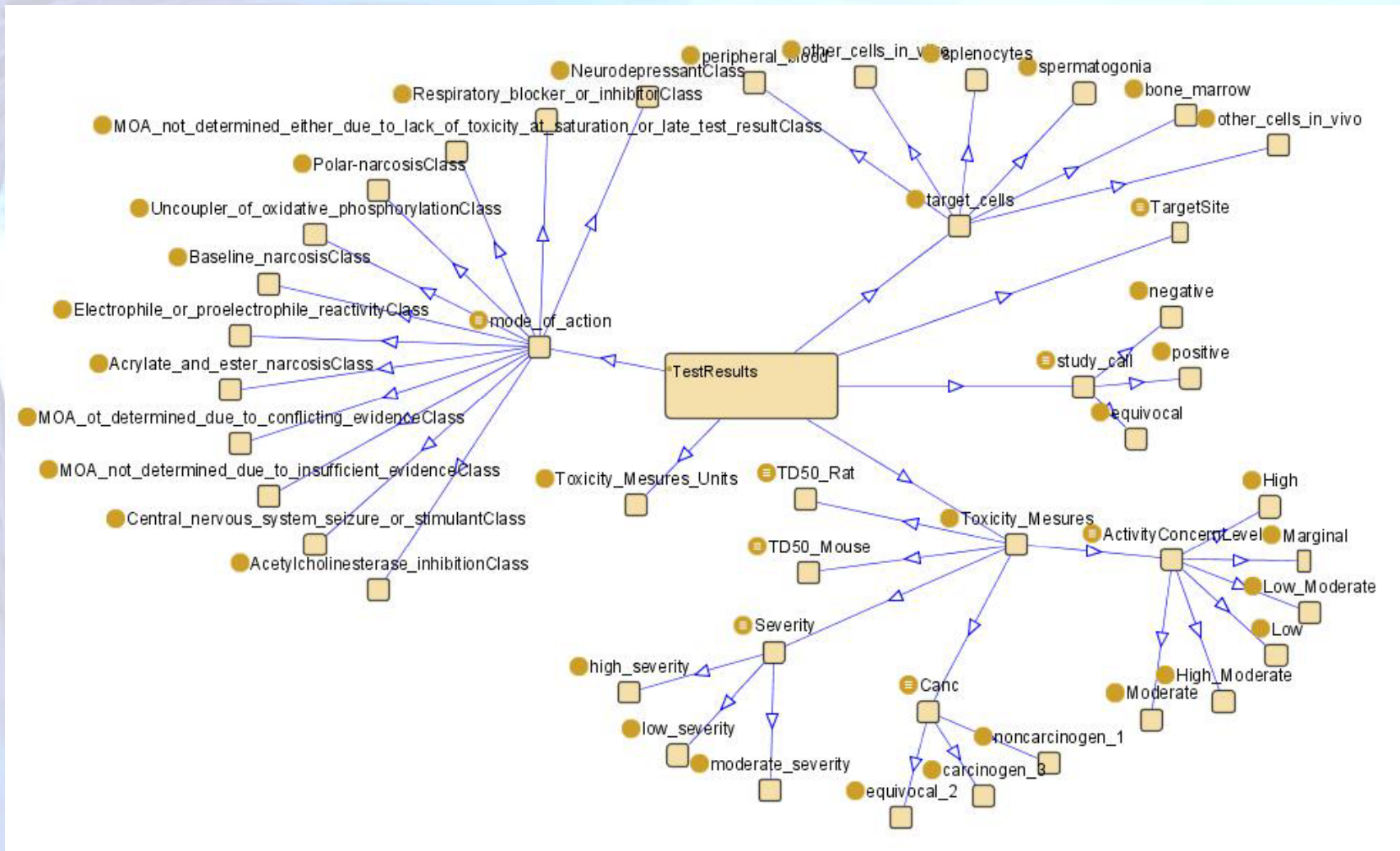
Main OWL Classes: TestConditions and TestResults



- Introduction of relationship between different classes, restrictions: e.g., introducing the property “has_Test_Species” limits 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)

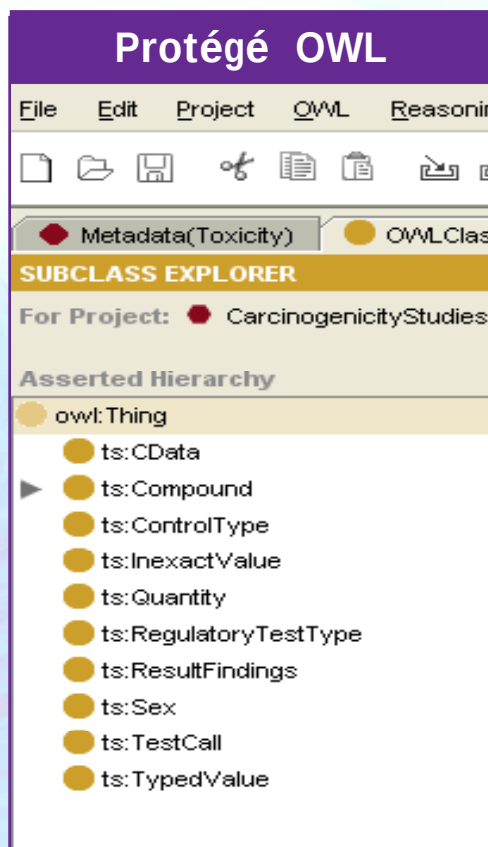
**ToxML
xsd
schema**



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

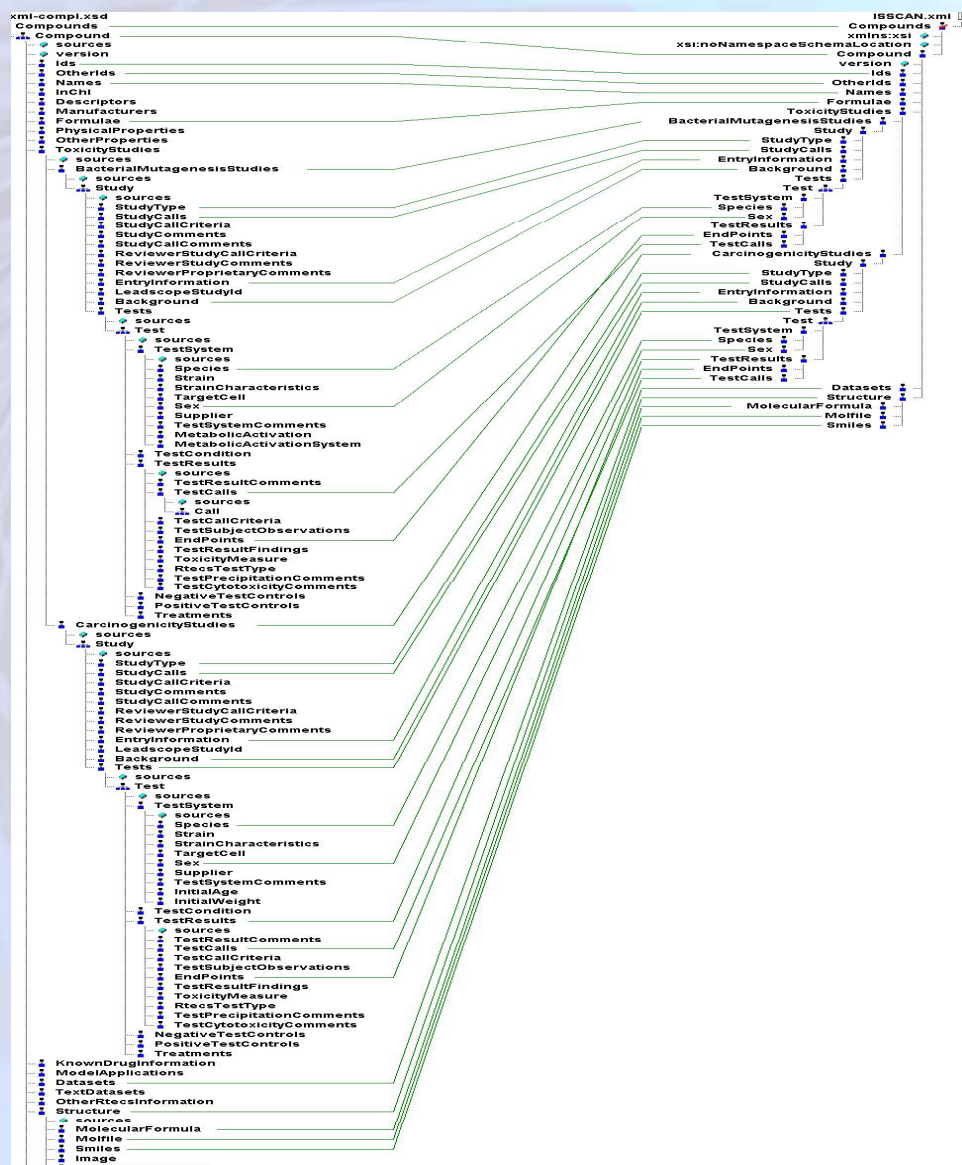
the parent relationships
are kept



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 Formyl trichloride; methane trichloride; methenyl trichloride; Methyl trichloride;
Synonyms	R 20; r 20 (refrigerant); Refrigerant R20; trichloroform; Trichloromethane
SAL	negative
Canc Reference	positive 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

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 and Diagnostic Criteria for Lesions in Rats and Mice)

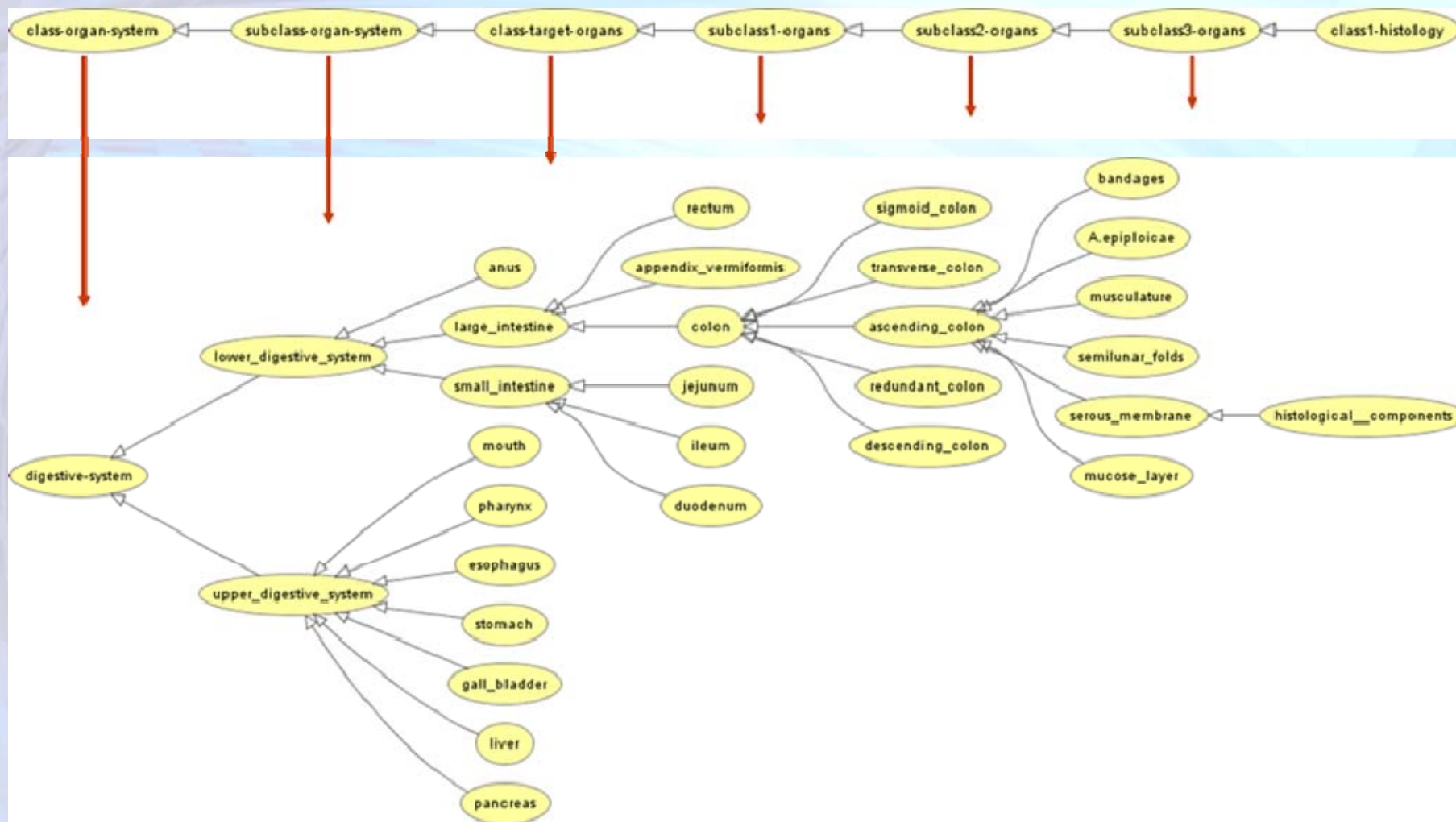
Organs system (class) –Subclass organs system

Target organs (class) - Targets organs (subclass 1 to N)

Histopathology (class) -Histopathology (subclasses if needed)

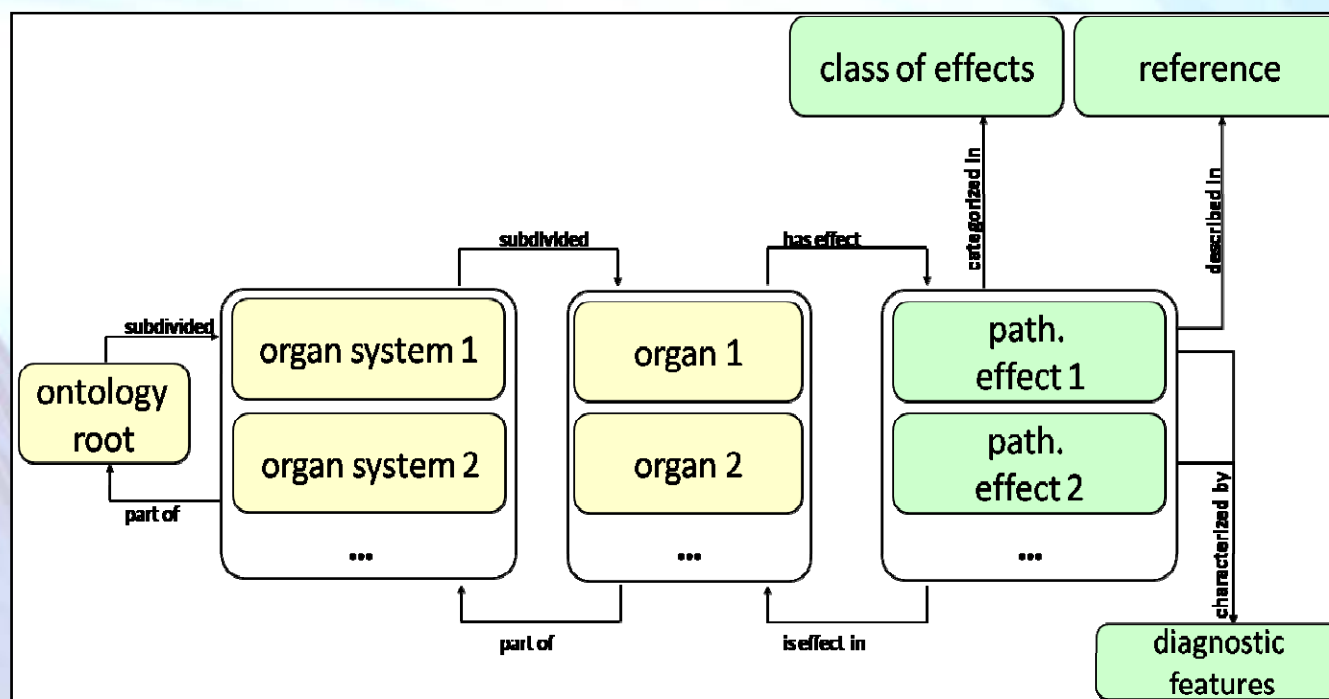
- **Key Features of the ITEM Organs Ontology:**
 - more synonyms regarding organ systems, target organs and their subclasses
 - organs are more detailed, up to histological components
 - organs linked to different organ systems (e.g. ovaries)
 - reviewed by pathologists, who have been involved in the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

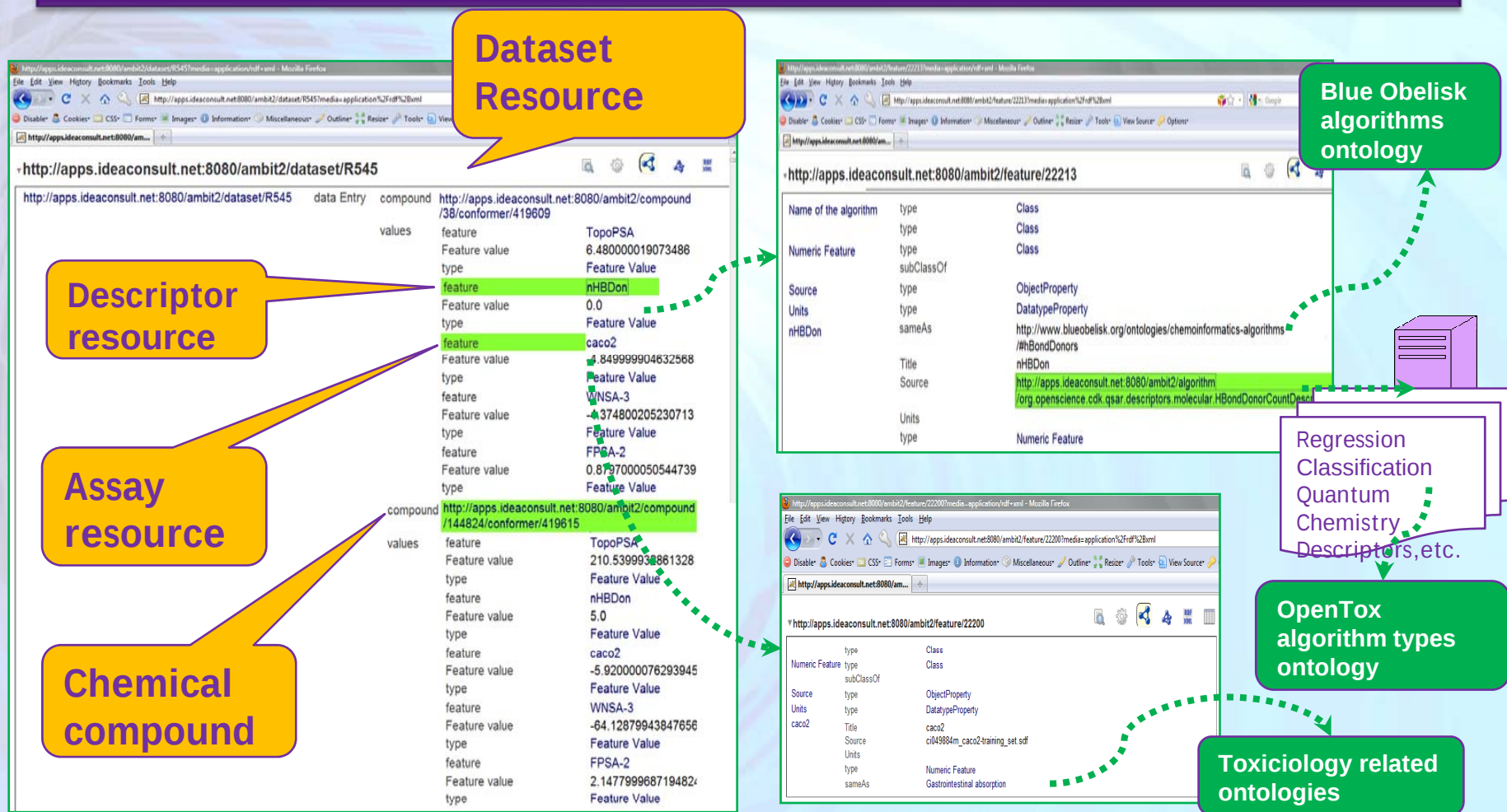


Structure of the combined Organs and Effects ontologies

The effects ontology consists of three main parts: classes of effects, linked to pathological effects, which are further linked 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.

The screenshot displays the ToxPredict web application interface, which is a demo application for estimating toxicological properties. The interface is divided into several sections:

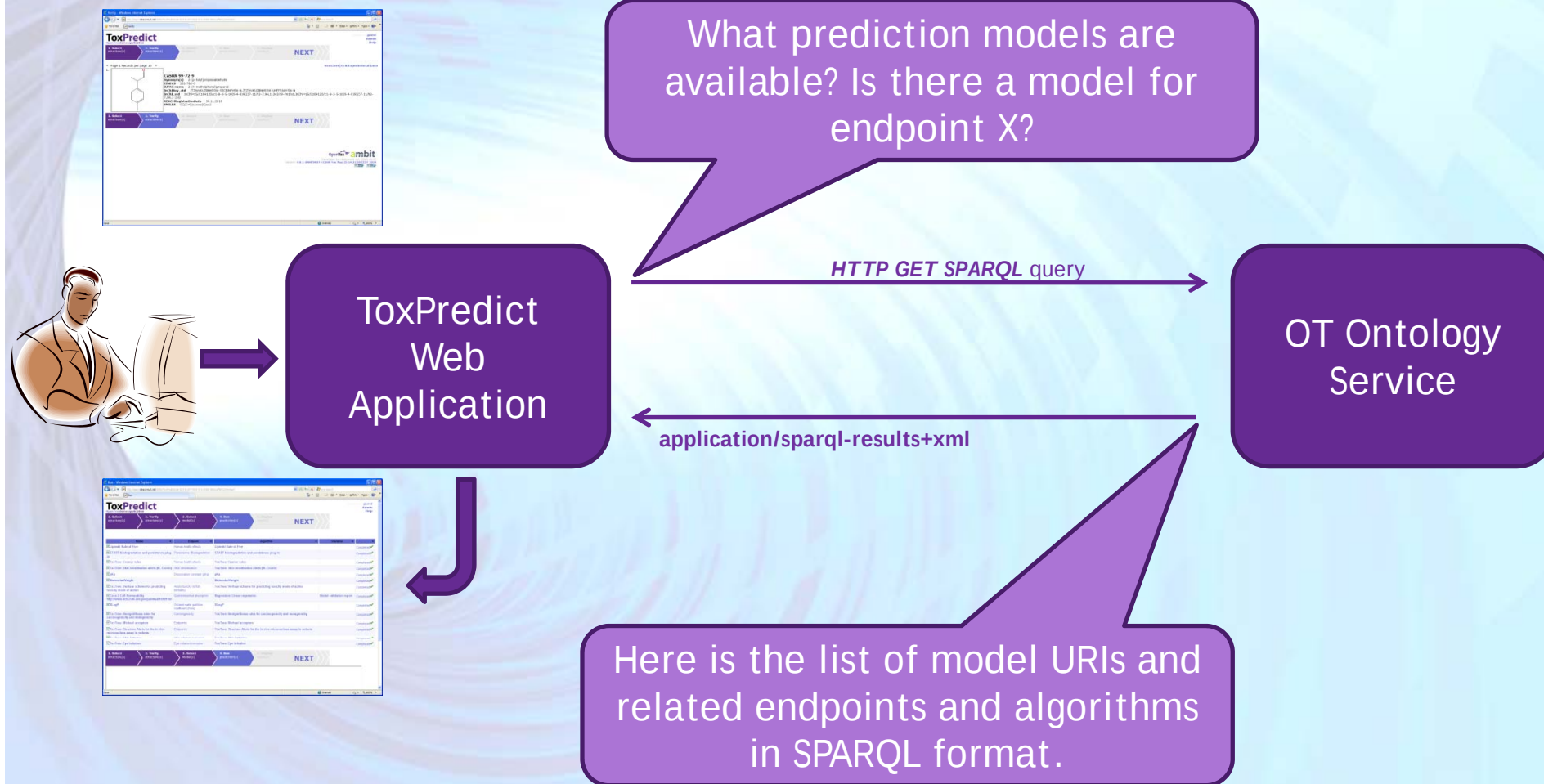
- Welcome - Mozilla Firefox**: The top browser window shows the ToxPredict logo and a brief description: "ToxPredict estimates the chemical hazard. Users can either search the OpenTox protocols (structures), grouped in 614 datasets, or use models, addressing 15 different endpoints models. Press NEXT to enter your ToxPredict."
- Verify - Mozilla Firefox**: The second browser window shows the "Verify" step of the workflow. It includes a chemical structure input field and a "Verify" button.
- Select - Mozilla Firefox**: The third browser window shows the "Select" step. It features a table of models and endpoints, and a "Select" button.
- Display - Mozilla Firefox**: The fourth browser window shows the "Display" step. It displays the results of the prediction, including the chemical structure, CASRN, and various toxicological endpoints.

The **Select** window contains a table of models and endpoints:

Model	Endpoint	Algorithm
☒ MolecularWeight		MolecularWeight
☐ ToxTree: Verhaar scheme for predicting toxicity mode of action	Acute toxicity to fish (lethality)	ToxTree: Verhaar
☐ ToxTree: Benigni/Bossa rules for carcinogenicity and mutagenicity	Carcinogenicity	ToxTree: Benigni/Bossa
☒ pKa	Dissociation constant (pKa)	pKa
☒ ToxTree: Structure Alerts for the in vivo micronucleus assay in rodents	Endpoints	ToxTree: Structure Alerts
☐ ToxTree: Michael acceptors	Endpoints	ToxTree: Michael acceptors
☐ ToxTree: Eye Irritation	Eye irritation/corrosion	ToxTree: Eye Irritation
☐ Caco-2 Cell Permeability	Gastrointestinal absorption	ToxTree: Caco-2 Cell Permeability
☐ InTox model created with	Gastrointestinal absorption	InTox model created with

The **Display** window shows the results for the selected model (MolecularWeight) and endpoint (pKa). The results include the chemical structure, CASRN (108-95-2), and various toxicological endpoints (LipinskiFailures, Dissociation_constant_pKa, Endpoints, HumanHealthEffects).

ToxPredict and Ontology Service (behind the scenes)

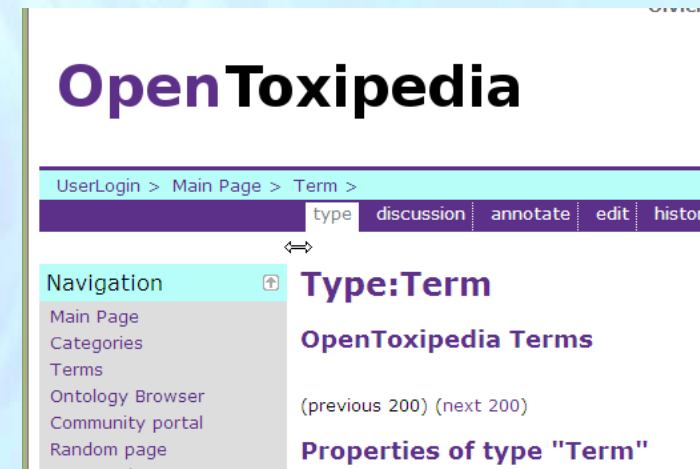


OpenToxipedia: community-based, predictive toxicology knowledge resource

- 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 annotation
 - 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 screenshot shows the OpenToxipedia website. At the top, there's a 'Log in' link. Below it, the 'OpenToxipedia' logo is on the left, and a search bar with 'Go' and 'Search' buttons is on the right. A breadcrumb trail reads 'Main Page > Categories >'. Below this is a navigation bar with tabs: 'category' (selected), 'discussion', 'view source', 'annotate', and 'history'. A left sidebar contains links for 'Navigation' (Main Page, Categories, Terms, Ontology Browser, Community portal, Random page, Recent changes, Help) and 'Toolbox' (What links here, Related changes, Special pages, Printable version, Permanent link, Browse properties). The main content area is titled 'Category:Categories' and 'OpenToxipedia Categories'. It includes instructions: 'to add a new term select a Category and click on Add or edit terms in this category use the form ... under Category headline.' Below this, it says 'Subcategories' and 'This category has the following 26 subcategories, out of 26 total.' A grid of subcategories is displayed, organized alphabetically from A to R.

A	D	I
• ADME • Acute • Adverse • Alternatives	• Developmental • Dose • Drug Safety	• In vitro
B	E	M
• Biology	• Ecotox • Endpoints • Epidemiology	• Mechanism • Mutagenicity
C	G	Q
• Carcinogenicity • Chemistry • Chronic	• Genotox	• QSAR
H	R	
		• Regulatory • Reproductive

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 Jeliaskova,
 - Vedrin Jeliaskov,
 - 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!