

# OpenTox Services and Applications

Barry Hardy (Douglas Connect)

AXLR8 Meeting  
Berlin, Germany  
23 May 2011

# Introduction

OpenTox has developed a novel semantic interoperable computing architecture for the field of predictive toxicology.

Today I will review what OpenTox services and applications can do.

# OpenTox is an Integrating Framework

## Framework

- Toxicity Data (Linked)
- *in silico* models
- Validation & Reporting
- Interpretation aids

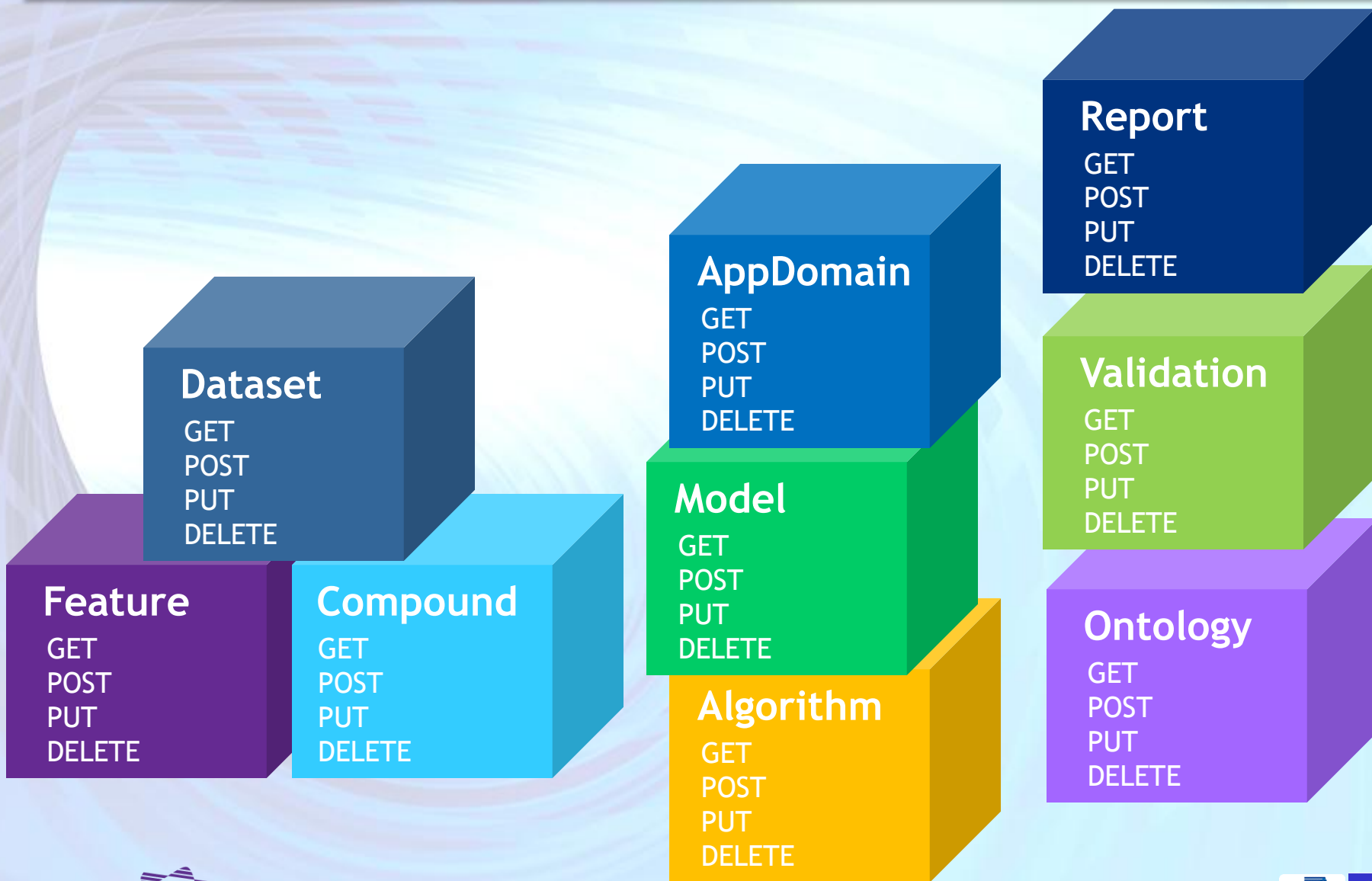
## Diverse Access

- Toxicologist, Biologist, Chemists
- Computational Scientists
- Interfaces for new algorithm development & integration

## Interoperability

- Promote Standards
- Core Open Source Components
- Support Ontologies & Integration of Multiple Resources

# Overview of OpenTox Application Programming Interfaces



# The OpenTox Framework (reported last year)

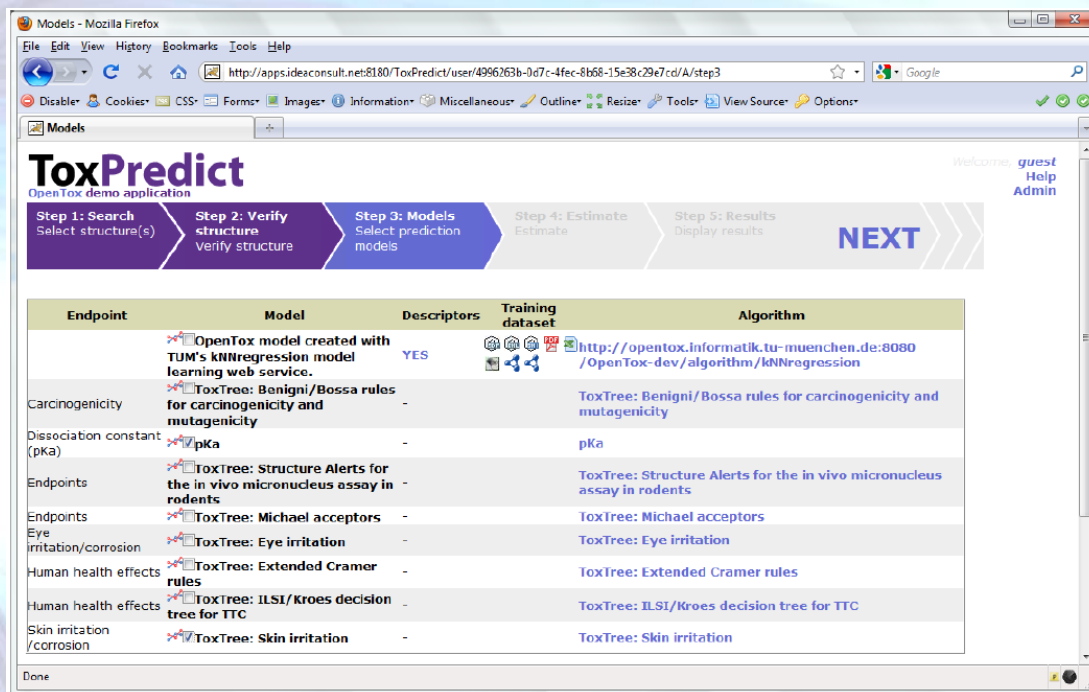
## Collaborative development of predictive toxicology applications


Journal of Cheminformatics 2010, 2:7 doi:10.1186/1758-2946-2-7

Barry Hardy, Nicki Douglas, Christoph Helma, Micha Rautenberg, Nina Jeliaskova, Vedrin Jeliaskov, Ivelina Nikolova, Romualdo Benigni, OlgaTcheremenskaia, Stefan Kramer, Tobias Girschick, Fabian Buchwald, Joerg Wicker, Andreas Karwath, Martin Gutlein, Andreas Maunz, Haralambos Sarimveis, Georgia Melagraki, Antreas Afantitis, Pantelis Sopasakis, David Gallagher, Vladimir Poroikov, Dmitry Filimonov, Alexey Zakharov, Alexey Lagunin, Tatyana Gloriovova, Sergey Novikov, Natalia Skvortsova, Dmitry Druzhilovsky, Sunil Chawla, Indira Ghosh, Surajit Ray, Hitesh Patel and Sylvia Escher

Open Access publication available at  
[www.jcheminf.com/content/2/1/7](http://www.jcheminf.com/content/2/1/7)

# What you can do with it ...

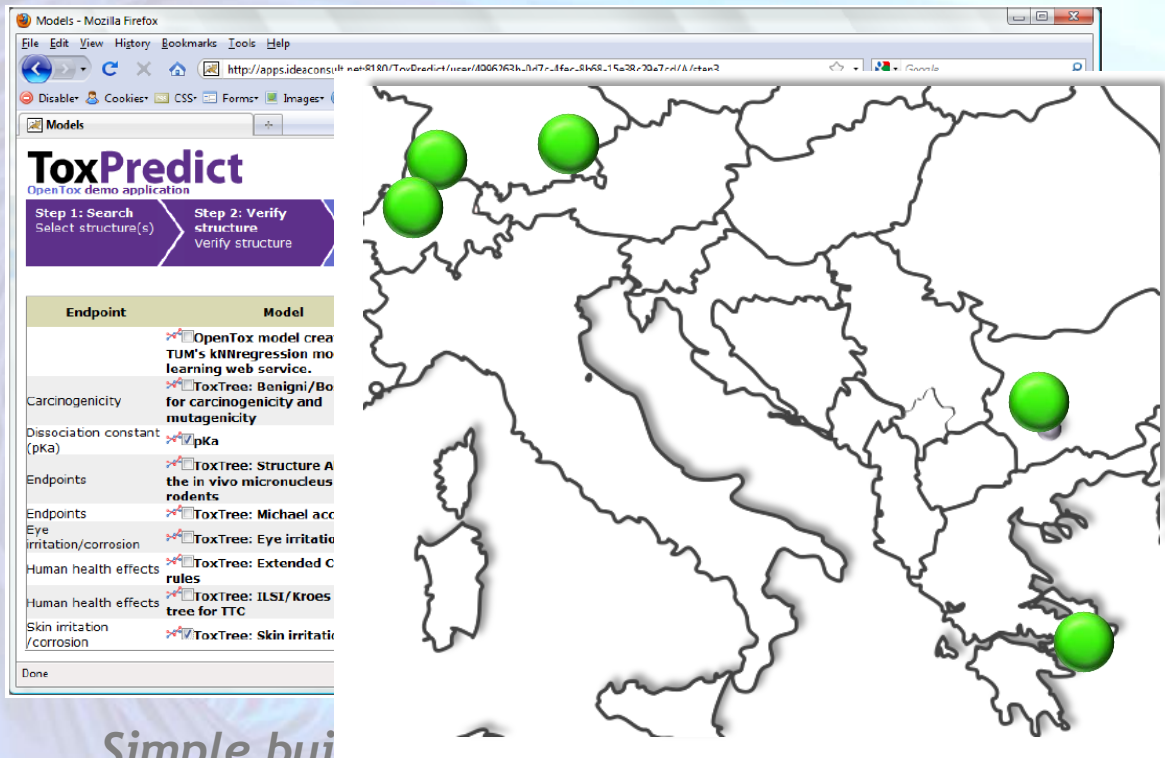


Endpoint	Model	Descriptors	Training dataset	Algorithm
	<input checked="" type="checkbox"/> OpenTox model created with TUM's kNNregression model learning web service.	YES	 <a href="http://opentox.informatik.tu-muenchen.de:8080/OpenTox-dev/algorithm/kNNregression">http://opentox.informatik.tu-muenchen.de:8080/OpenTox-dev/algorithm/kNNregression</a>	
Carcinogenicity	<input checked="" type="checkbox"/> ToxTree: Benigni/Bossa rules for carcinogenicity and mutagenicity	-		ToxTree: Benigni/Bossa rules for carcinogenicity and mutagenicity
Dissociation constant (pKa)	<input checked="" type="checkbox"/> pKa	-		pKa
Endpoints	<input checked="" type="checkbox"/> ToxTree: Structure Alerts for the in vivo micronucleus assay in rodents	-		ToxTree: Structure Alerts for the in vivo micronucleus assay in rodents
Endpoints	<input checked="" type="checkbox"/> ToxTree: Michael acceptors	-		ToxTree: Michael acceptors
Eye irritation/corrosion	<input checked="" type="checkbox"/> ToxTree: Eye irritation	-		ToxTree: Eye irritation
Human health effects	<input checked="" type="checkbox"/> ToxTree: Extended Cramer rules	-		ToxTree: Extended Cramer rules
Human health effects	<input checked="" type="checkbox"/> ToxTree: ILSI/Kroes decision tree for TTC	-		ToxTree: ILSI/Kroes decision tree for TTC
Skin irritation/corrosion	<input checked="" type="checkbox"/> ToxTree: Skin irritation	-		ToxTree: Skin irritation

*Simple building of predictive toxicology applications based on well-established methods and databases*



# What you can do with it ...



The screenshot shows the ToxPredict web application in a Mozilla Firefox browser. The interface includes a navigation bar with 'Step 1: Search' and 'Step 2: Verify'. Below this is a table listing various endpoints and the models used for prediction.

Endpoint	Model
Carcinogenicity	OpenTox model creation TUM's kNN regression machine learning web service. ToxTree: Benigni/Bo for carcinogenicity and mutagenicity
Dissociation constant (pKa)	pKa
Endpoints	ToxTree: Structure A the in vivo micronucleus rodents
Endpoints	ToxTree: Michael acc
Eye irritation/corrosion	ToxTree: Eye irritation
Human health effects	ToxTree: Extended C rules
Human health effects	ToxTree: ILSI/Kroes tree for TTC
Skin irritation/corrosion	ToxTree: Skin irritation

Below the table, there is a 'Done' button. To the right of the browser window, a map of Europe is displayed with five green circular markers indicating specific locations: three in the northwestern part of Europe (Ireland, United Kingdom, France) and two in the southeastern part (Greece, Turkey).

Simple built-in  
application  
methods and

*Distributed applications,  
integrating wide range of data,  
models, prediction methods*

# What you can do with it ...

The image displays two software interfaces. On the left is the 'ToxPredict OpenTox demo application' in a Mozilla Firefox browser. It features a map of Europe with three green dots and a table of endpoints and models. On the right is the 'Taverna Workbench 2.1.0' interface, showing a workflow diagram with steps like 'ask\_username', 'choose\_trainset', 'upload\_trainset', 'calculate\_descriptors', 'learn\_model', and 'apply\_model\_to\_testset'.

**ToxPredict OpenTox demo application**

Step 1: Search Select structure(s) Step 2: Verify structure Verify structure

Endpoint	Model
Carcinogenicity	OpenTox model creat TUM's kNNregression mod learning web service.
Dissociation constant (pKa)	ToxTree: Benigni/Bos for carcinogenicity and mutagenicity
Endpoints	ToxTree: Structure Al the in vivo micronucleus rodents
Endpoints	ToxTree: Michael acc
Eye irritation/corrosion	ToxTree: Eye irritation rules
Human health effects	ToxTree: Extended Cr
Human health effects	ToxTree: ILSI/Kroes c tree for TTC
Skin irritation /corrosion	ToxTree: Skin irritation

Done

**Taverna Workbench 2.1.0**

Workflow diagram

Workflow input ports: message\_value\_2, title\_value\_4

Workflow output ports: result

Workflow steps: ask\_username, choose\_trainset, upload\_trainset, calculate\_descriptors, learn\_model, apply\_model\_to\_testset, wait\_for\_prediction

Simple building blocks for applications  
Distributed applications  
methods and a wide range of methods

Integration into workflow systems  
for computational biology



# Need for Applications for REACH



# REACH and (Q)SAR bottlenecks

Wim De Coen, ECHA, “*Current Challenges from Evaluation Point of View - Introduction Case Studies*”, ECHA Experts Workshop on “Dealing with Uncertainty of Non-Test Methods under REACH”, 2010:

## Specific Bottlenecks for (Q)SAR:

- Well standardized and accepted OECD principles
- Issues mainly at level of documentation
  - Level of documentation insufficient
    - QMRF, QPRF missing
  - Applicability domain unclear
  - Unclear training datasets & algorithm
    - General issue of well established/commercial QSAR packages

# REACH and data bottlenecks

There exists considerable uncertainty in decision making based on current reproductive toxicity data, which place the largest potential demands on animal testing required by REACH.

Improvements to reduce uncertainty in decision making require:  
a “robust reference dataset of harmonised test information”

Reference: Dick Sijm and Betty Hakkert, RIVM, *“Use of non-test methods in integrated testing strategies for making informed decisions - Non-test methods require robust reference datasets”*, ECHA Experts Workshop on “Dealing with Uncertainty of Non-Test Methods under REACH”, 2010)



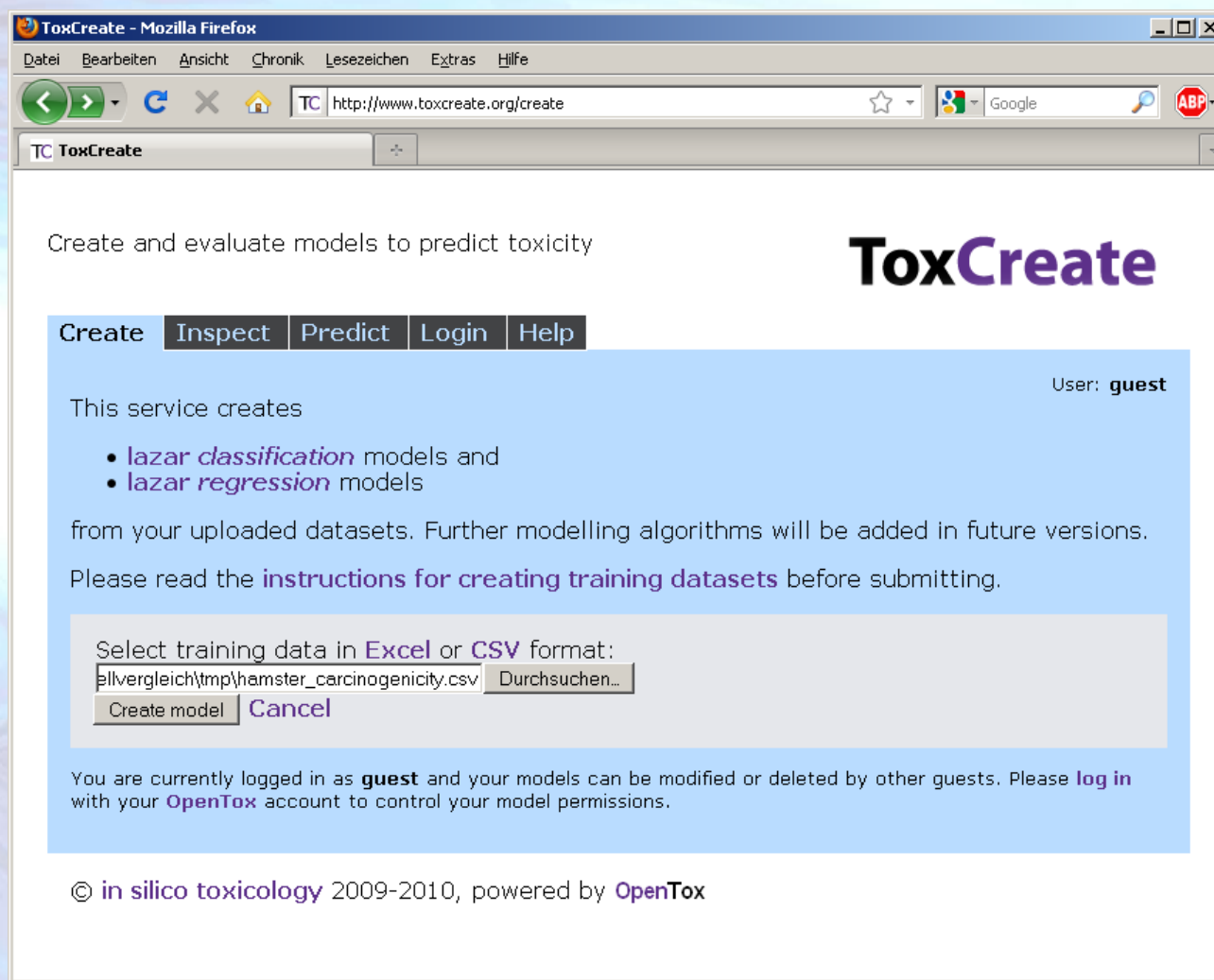
# (Q)SARs & REACH requirements

(Quantitative) Structure Activity Relationship = (Q)SAR

According to REACH Annex XI, (Q)SAR results may be used instead of testing when all of the following conditions are met:

- The results are derived from a (Q)SAR model whose **scientific validity** has been established.
- The substance falls within the **applicability domain** of the (Q)SAR model.
- The results are **adequate** for the purpose of classification and labeling and/or risk assessment.
- Adequate and reliable **documentation** of the applied method is provided.

# ToxCreate - (Q)SAR Model Building application



The screenshot shows the ToxCreate web application running in a Mozilla Firefox browser. The browser's address bar displays the URL <http://www.toxcreate.org/create>. The page title is "ToxCreate - Mozilla Firefox". The application interface includes a navigation menu with links: "Datei", "Bearbeiten", "Ansicht", "Chronik", "Lesezeichen", "Extras", and "Hilfe". Below the navigation menu, there is a search bar with the text "TC ToxCreate" and a search button. The main content area features the heading "Create and evaluate models to predict toxicity" and the "ToxCreate" logo. A navigation bar contains buttons for "Create", "Inspect", "Predict", "Login", and "Help". The "Create" button is highlighted. The main content area states: "This service creates" followed by a list of model types: "lazar classification" and "lazar regression" models. It also mentions that further modelling algorithms will be added in future versions. A note advises users to read the "instructions for creating training datasets" before submitting. A form for selecting training data is shown, with a text input field containing the file path "e:\vergleich\tmp\hamster\_carcinogenicity.csv" and a "Durchsuchen..." button. Below the input field are "Create model" and "Cancel" buttons. The page also displays the user status "User: guest" and a message indicating that the user is currently logged in as "guest" and their models can be modified or deleted by other guests, with a prompt to "log in" with their "OpenTox" account to control model permissions. The footer contains the copyright notice: "© in silico toxicology 2009-2010, powered by OpenTox".

Create and evaluate models to predict toxicity

## ToxCreate

Create Inspect Predict Login Help

User: **guest**

This service creates

- **lazar classification** models and
- **lazar regression** models

from your uploaded datasets. Further modelling algorithms will be added in future versions.

Please read the **instructions for creating training datasets** before submitting.

Select training data in **Excel** or **CSV** format:

You are currently logged in as **guest** and your models can be modified or deleted by other guests. Please **log in** with your **OpenTox** account to control your model permissions.

© in silico toxicology 2009-2010, powered by **OpenTox**

Developed by In Silico Toxicology



# ToxCreat - (Q)SAR Model Results

ToxCreat - Mozilla Firefox

File Bearbeiten Ansicht Chronik Lesezeichen Extras Hilfe

TC http://www.toxcreate.org/models

TC ToxCreat

Create and evaluate models to predict toxicity

## ToxCreat

Create Inspect Predict Login Help

User: guest

Get an overview about ToxCreat models. Parts of this page are refreshed every 5 seconds to update the model status.

### Hamster Carcinogenicity ( edit )

Status: Completed(delete)

Training compounds: 85

Algorithm: lazar

Type: classification

Descriptors: Fminer backbone refinement classes

Training dataset: Excel sheet , YAML (experts)

Feature dataset: Excel sheet , YAML (experts)

Model: QMRF Editor, YAML (experts, models cannot be represented in Excel)

Validation: show

Detailed report:

Number of predictions: 69

Correct predictions: 82.68 %

Weighted area under ROC: 0.935

Specificity: 0.143

Sensitivity: 0.865

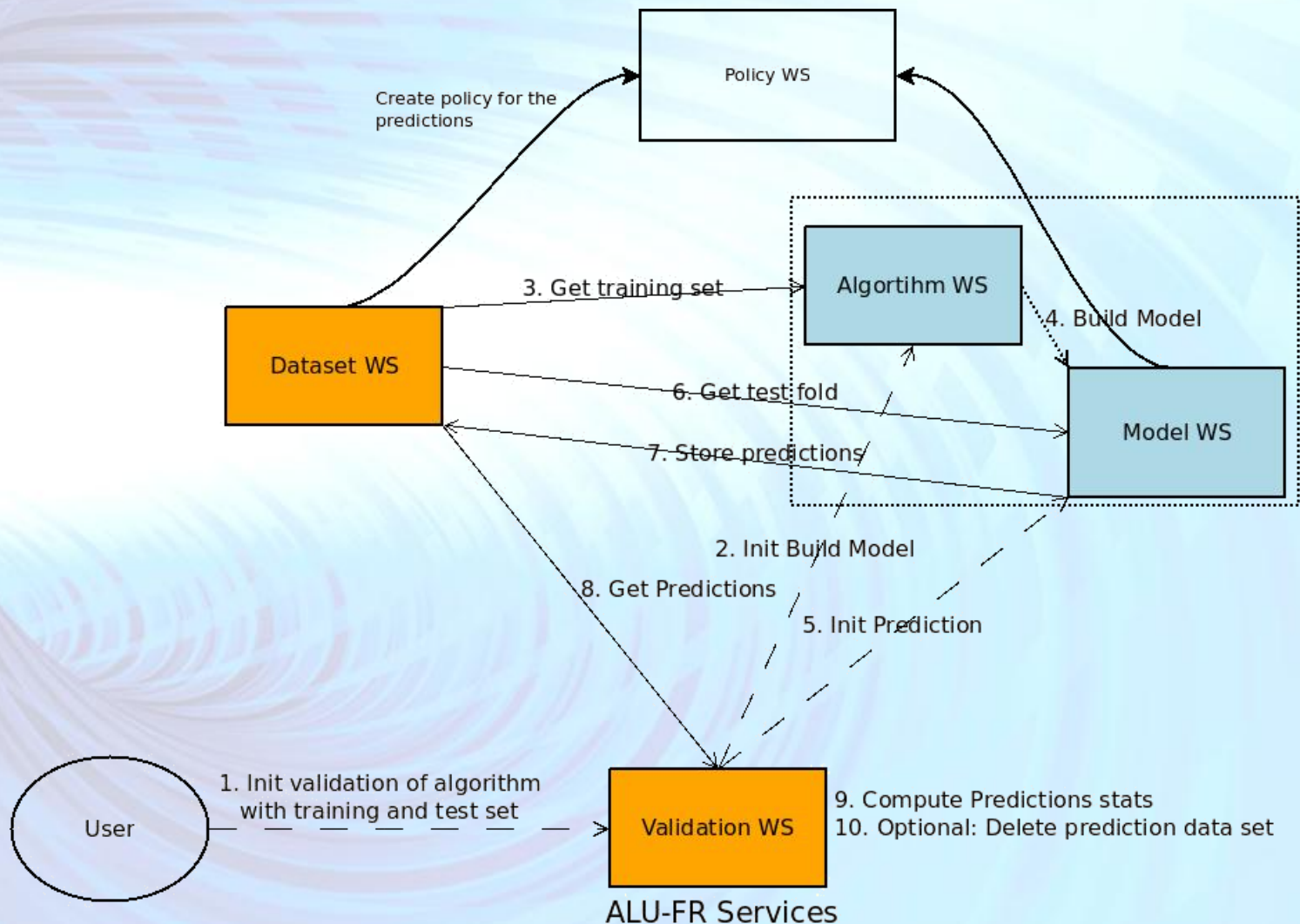
Confusion Matrix:

		Measured	
		active	inactive
Predicted	active	32	5
	inactive	7	25

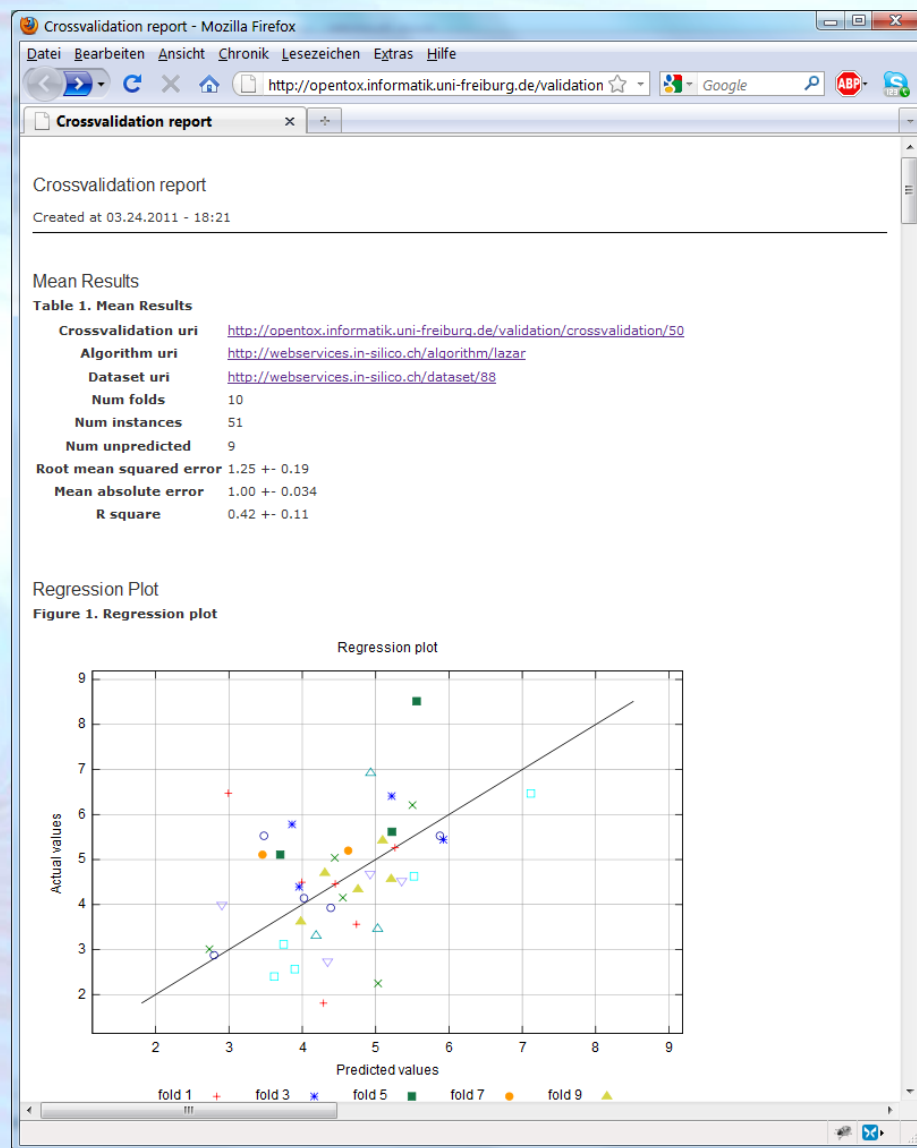
© in silico toxicology 2009-2010, powered by OpenTox

	OECD Principle	OpenTox addresses Validation Principles by...
1	Defined Endpoint	providing a unified source of well defined and documented toxicity data with a common vocabulary
2	Unambiguous Algorithm	providing transparent access to well documented models and algorithms as well as to the source code
3	Defined Applicability Domain	integrating tools for the determination of applicability domains during the validation of prediction models
4	Goodness-of-fit, robustness and predictivity	providing scientifically sound validation routines for the determination of errors and confidences
5	Mechanistic interpretation (if possible)	integrating tools for the inference, correlation or prediction of toxicological mechanisms and the recording of opinions and analysis in reports

# Validation within OpenTox



# ToxCreate - linked to Validation Service



# ToxCreate - Confidence, Supporting Information

ToxCreate - Mozilla Firefox

Datei Bearbeiten Ansicht Chronik Lesezeichen Extras Hilfe

TC http://www.toxcreate.org/lazar#lazar\_algorithm

TC ToxCreate phenylhydrazine (CHEBI:27924)

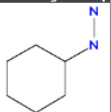
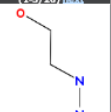
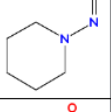
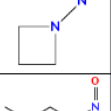
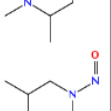
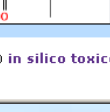
Create and evaluate models to predict toxicity

**ToxCreate**

Create Inspect **Predict** Login Help

User: guest

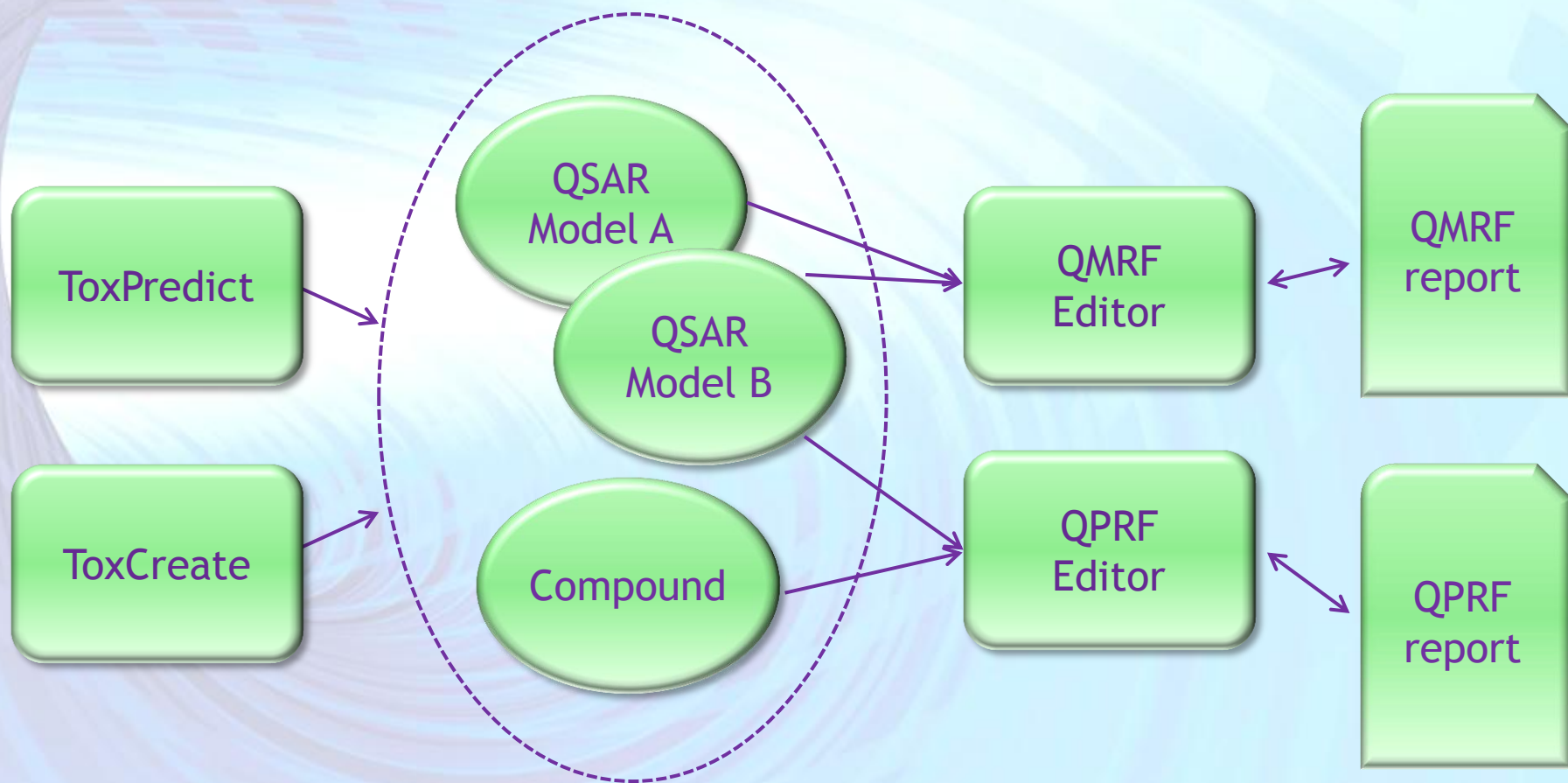
New prediction

Hamster Carcinogenicity	Prediction	Confidence	Supporting information
	active	0.108	Names and synonyms Significant fragments
Neighbors (1-5/25) next	Measured activity	Similarity	Supporting information
	inactive	0.715	Names and synonyms Significant fragments
	inactive	0.5	Names and synonyms Significant fragments
	inactive	0.5	Names and synonyms Significant fragments
	inactive	0.5	Names and synonyms Significant fragments
	inactive	0.5	Names and synonyms Significant fragments

© in silico toxicology 2009-2010, powered by OpenTox



# (Q)SARs - reporting in OpenTox



# (Q)SARs - QMRF reporting in OpenTox

QMRF Editor 0.05 OpenTox Version [http://opentox.informatik.uni-freiburg.de/validation/reach\\_report/QMRF/3](http://opentox.informatik.uni-freiburg.de/validation/reach_report/QMRF/3)

File Edit Style

**QMRF (Q)SAR Model Reporting Format (QMRF), Version 1.2**

Welcome

Version

Name

Section 1. Author

Section 2. Date

Section 3. Contact

Section 4. Email

Section 5. www

Section 6. **Background**

Section 7. The set of information that you provide will be used to generate (Q)SARs. For this purpose, the structure of the QMRF is devised to reflect as much as possible the OECD principles for the validation, for regulatory purposes, of (Q)SAR models.

Section 8. You are invited to consult the OECD "Guidance Document on the Validation of (Quantitative) Structure-Activity Relationship Models" that can aid you in filling in a number of fields of the QMRF (visit the following webpage for downloading the proper documentation: [http://ecb.jrc.it/qsar/background/background\\_oecd\\_principles.php](http://ecb.jrc.it/qsar/background/background_oecd_principles.php))

Section 9. **Submission Procedure**

Section 10. If you wish to submit the QMRF for inclusion in the JRC QSAR Model Database, please save your QMRF as xml file and upload it by the on-line submission procedure

Download started

Eingabe

Please enter the URI for the download

OK Abbrechen

# QPRF Reporting (Qedit)

Qedit application window showing the Applicability Domain Info and Compound Details sections.

**3.3. Applicability Domain Info.**

Name Applicability Domain Estimation Algorithm Used :

Link to Applicability Domain Resource :

**3.3.b. Structural Analogues**

Add Compound Wizard Remove Clear List Similarity Level:  Acquire List of Analogues Compound Info

List of Structural Analogues (URIs):

Chemical Name	Experimental Value
phenobarbital, Phen...	
5-methyl-5-phenylb...	
methylphenobarbit...	
5-allyl-5-phenylbar...	
primidone, Primaclo...	
calcium bis[5-(1-cyc...	
5-ethyl-5-(4'-hydrox...	
barbexaclone	
1,3-dimethyl-5-phen...	
5-ethyl-5-phenylbar...	
N-(acetaminophenyl)...	

Image of structural analogue:

**3.3.c. Consideration**

**Discussion**

Applicability Domain Result: ☒

3.3.a. Choose Domain :

**Compound Details**

URI:

Smiles:

InChI:

InChI Key:

CAS number:

Chemical Name:

Einecs:

REACH Reg. Date:

Available Conformers (Links):

- <http://ambit.uni-plovdiv.bg:8080/ambit2/compound/5100/conformer/5100>
- <http://ambit.uni-plovdiv.bg:8080/ambit2/compound/5100/conformer/105301>
- <http://ambit.uni-plovdiv.bg:8080/ambit2/compound/5100/conformer/181274>

Close Apply Changes and Close

Application by Pantelis Sopasakis (NTUA)

# Metabolites

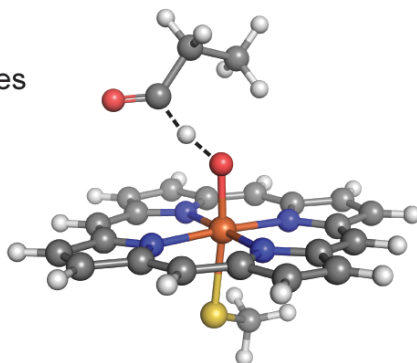
According to ECHA Guidance B, further investigation may be required for degradation products and metabolites if considered relevant for the chemical safety assessment, PBT assessment or classification and labeling.

# SMARTCyp Service for Predicting Metabolites

## Atom Reactivity Library

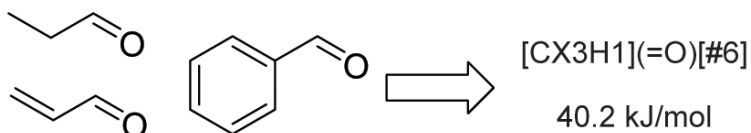
### A. Calculate Quantum Chemical Reference Energies

Calculate transition state energies using density functional theory



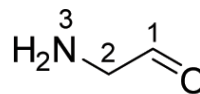
### B. Define SMARTS Rules

Group calculations by fragments and calculate average energies



## SMARTCyp

### 1. Assign Energies By SMARTS matching



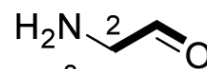
Atom	SMARTS	Energy
1	<chem>[CX3H1](=O)[#6]</chem>	40.2
2	<chem>[CX4][N]</chem>	39.8
3	<chem>[N^3][H1,H2]</chem>	54.1

### 2. Compute Accessibility Descriptor

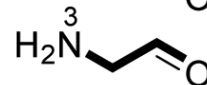
$$A_i = \text{Maxbonds}_i / \text{Maxbonds}_{\text{all}}$$



$$A_1 = 2 / 3 = 0.67$$



$$A_2 = 2 / 3 = 0.67$$



$$A_3 = 3 / 3 = 1.00$$

### 3. Compute Score and Rank Atoms

$$\text{Score, } S = E - 8A$$

Lowest score gets rank 1

$$S_1 = 40.2 - 8 \cdot 0.67 = 34.84$$

Atom 1 - Rank 2

$$S_2 = 39.8 - 8 \cdot 0.67 = 34.44$$

Atom 2 - Rank 1

$$S_3 = 54.1 - 8 \cdot 1.00 = 46.10$$

Atom 3 - Rank 3



SMARTCyp - developed by Patrik Rydberg, University of Copenhagen

[www.farma.ku.dk/index.php/SMARTCyp/7990/0/](http://www.farma.ku.dk/index.php/SMARTCyp/7990/0/)



# Metabolites

[ToxPredict](#) [TTC](#) [Depiction](#) [Datasets](#) [Chemical compounds](#) [Similarity](#) [Substructure](#)  
[playground](#) [Help](#)

# ambit

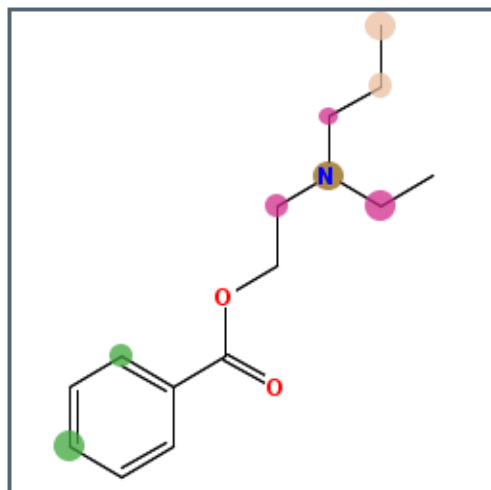
Search by property or identifier name  
(optional) and value  
*This site and AMBIT REST services  
are under development!*

Sea



idconformer=773441

Save search results



-  Aliphatic Hydroxylation
-  Aromatic Hydroxylation
-  N-Dealkylation
-  N-Oxidation
-  O-Dealkylation
-  S-Oxidation

Developed by Ideaconsult

# Information Gathering & Evaluation Process

Annex VI of REACH describes four steps of the Information Gathering and Evaluation Process (IGEP) that need to be followed by the registrant to fulfill the information requirements for a substance:

**Step 1** Gather and share existing information

**Step 2** Consider information needs

**Step 3** Identify information gaps

**Step 4** Generate new information or propose a testing strategy

# Information Gathering & Evaluation Process

In step 1 of the IGEP, the registrant must collect all existing *in vivo* toxicological and ecotoxicological information that is relevant and available to him regardless of whether information on a given endpoint is required or not at the specific tonnage level.

The Part B guidance states that this information may be obtained from a variety of sources including “databases or other sources in the open literature or accessible on the internet”.

**Step 1: Search**  
Select structure(s)

**Step 2: Verify  
structure**  
Verify structure

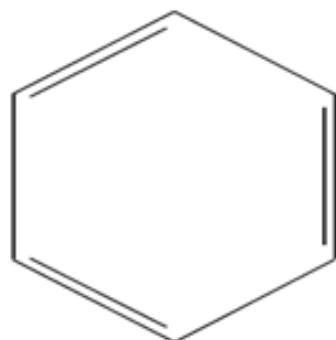
**Step 3: Models**  
Select prediction  
models

**Step 4: Estimate**  
Estimate

**Step 5: Results**  
Display results

This page lists your ToxPredict workflow results for the structure(s) you have selected and the model prediction(s) you have chosen to run. You could also retrieve the ToxPredict report in various other formats, e.g. [SDF](#), [CML](#), [SMI](#), [PDF](#), [CSV](#), [ARFF](#), [RDF/XML](#) or [RDF/N3](#).

Download as 



**CAS RN**  
**EINECS**  
**IUPAC name**  
**Synonym**

71-43-2  
200-753-7  
benzene  
(6)annulene; benzine; Benzol; Benzolene;  
bicarburet of hydrogen; carbon oil; Coal naphtha;  
cyclohexatriene; mineral naphtha; motor benzol;  
nitration benzene; Phene; Phenyl hydride;  
pyrobenzol.

**Synonym**  
**Synonym**  
**Synonym**  
**Quality label**

21742.0  
Benzene  
benzene  
OK

**MolecularWeight**  **MolecularWeight**

**MW**

78.1112

Select
Welcome, [guest](#)  
[Admin](#)  
[Help](#)

# ToxPredict

OpenTox demo application

1. Select  
structure(s)

2. Verify  
structure(s)

**3. Select  
model(s)**

4. Run  
prediction(s)

5. Display  
result(s)

NEXT

Model	Endpoint	Algorithm	Validation
<input checked="" type="checkbox"/> MolecularWeight		MolecularWeight	
<input checked="" type="checkbox"/> ToxTree: Verhaar scheme for predicting toxicity mode of action	Acute toxicity to fish (lethality)	ToxTree: Verhaar scheme for predicting toxicity mode of action	
<input checked="" type="checkbox"/> ToxTree: Benigni/Bossa rules for carcinogenicity and mutagenicity	Carcinogenicity	ToxTree: Benigni/Bossa rules for carcinogenicity and mutagenicity	
<input checked="" type="checkbox"/> pKa	Dissociation constant (pKa)	pKa	
<input checked="" type="checkbox"/> ToxTree: Structure Alerts for the in vivo micronucleus assay in rodents	Endpoints	ToxTree: Structure Alerts for the in vivo micronucleus assay in rodents	
<input checked="" type="checkbox"/> ToxTree: Michael acceptors	Endpoints	ToxTree: Michael acceptors	
<input checked="" type="checkbox"/> ToxTree: Eye irritation	Eye irritation/corrosion	ToxTree: Eye irritation	
<input checked="" type="checkbox"/> Caco-2 Cell Permeability <a href="http://www.ncbi.nlm.nih.gov/pubmed/16959190">http://www.ncbi.nlm.nih.gov/pubmed/16959190</a>	Gastrointestinal absorption	Regression: Linear regression	Model validation report
<input checked="" type="checkbox"/> OpenTox model created with TUM's PLScression model learning web service.	Gastrointestinal absorption	<a href="http://opentox.informatik.tu-muenchen.de:8080/OpenTox-dev/algorithm/PLScression">http://opentox.informatik.tu-muenchen.de:8080/OpenTox-dev/algorithm/PLScression</a>	
<input checked="" type="checkbox"/> OpenTox model created with TUM's kNNression model learning web service.	Gastrointestinal absorption	<a href="http://opentox.informatik.tu-muenchen.de:8080/OpenTox-dev/algorithm/kNNression">http://opentox.informatik.tu-muenchen.de:8080/OpenTox-dev/algorithm/kNNression</a>	
<input checked="" type="checkbox"/> <a href="http://opentox.ntua.gr:3000/model/679a80e6-b2d9-45c1-ba42-4489e85b5898">http://opentox.ntua.gr:3000/model/679a80e6-b2d9-45c1-ba42-4489e85b5898</a>	Gastrointestinal absorption	Multiple Linear Regression Training Algorithm	
<input checked="" type="checkbox"/> Lipinski Rule of Five	Human health effects	Lipinski Rule of Five	
<input checked="" type="checkbox"/> ToxTree: Cramer rules	Human health effects	ToxTree: Cramer rules	
<input checked="" type="checkbox"/> XLogP	Octanol-water partition coefficient (Kow)	XLogP	
<input checked="" type="checkbox"/> START biodegradation and persistence plug-in	Persistence: Biodegradation	START biodegradation and persistence plug-in	
<input checked="" type="checkbox"/> SmartCYP: Cytochrome P450-Mediated Drug Metabolism	Protein-binding	SmartCYP: Cytochrome P450-Mediated Drug Metabolism	
<input checked="" type="checkbox"/> ToxTree: Skin irritation	Skin irritation /corrosion	ToxTree: Skin irritation	
<input checked="" type="checkbox"/> ToxTree: Skin sensitisation alerts (M. Cronin)	Skin sensitisation	ToxTree: Skin sensitisation alerts (M. Cronin)	



# Information Gathering & Evaluation Process

Display

## ToxPredict

OpenTox demo application

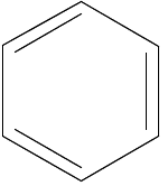
Welcome [guest](#) [Admin](#) [Help](#)

1. Select structure(s) 2. Verify structure(s) 3. Select model(s) 4. Run prediction(s) 5. Display result(s)

<< Page 1 Records per page 1 >>

Structure(s) & Model predictions & Experimental Data SDF CSV PDF

1.



**CASRN 71-43-2**  
**Synonym(s)** benzene,(6)annulene; benzine; Benzol; Benzolene; bicarburet of hydrogen; carbon oil; Coal naphtha; cyclohexatriene; mineral naphtha; motor benzol; nitration benzene; Phene; Phenyl hydride; pyrobenzol, Benzene, BENZENE, C6H6  
**EINECS** 200-753-7  
**IUPAC name** benzene  
**InChIKey\_std** UHOVQNZJYSORNB-UHFFFAOYSA-N  
**InChI\_std** InChI=1S/C6H6/c1-2-4-6-5-3-1/h1-6H  
**REACHRegistrationDate** 30.11.2010  
**SMILES** c1ccccc1,C1=CC=CC=C1

OpenTox model created with TUM's PLSRegression model learning web service.

Prediction feature for <a href="http://apps.ideaconsult.net:8080/ambit2/feature/22200">http://apps.ideaconsult.net:8080/ambit2/feature/22200</a> endpoint prediction	-4.496099948883057
<b>LipinskiFailures</b> <b>Lipinski Rule of Five</b>	
LipinskiFailures	0.0
<b>Acute_toxicity_to_fish_lethality</b> <b>ToxTree: Verhaar scheme for predicting toxicity mode of action</b> Verhaar scheme	Class 1 (narcosis or baseline toxicity)
<b>Carcinogenicity</b> <b>ToxTree: Benigni/Bossa rules for carcinogenicity and mutagenicity</b>	
Structural Alert for genotoxic carcinogenicity	NO
Structural Alert for nongenotoxic carcinogenicity	NO
Potential S. typhimurium TA100 mutagen based on QSAR	NO
Unlikely to be a S. typhimurium TA100 mutagen based on QSAR	NO
Potential carcinogen based on QSAR	NO
Unlikely to be a carcinogen based on QSAR	NO
For a better assessment a QSAR calculation could be applied.	NO
Negative for genotoxic carcinogenicity	YES
Negative for nongenotoxic carcinogenicity	YES
Structural Alert for genotoxic carcinogenicity#explanation	QA1.Acyl halides <b>No</b> QA2.Alkyl (C<5) or benzyl ester of sulphonic or phosphonic acid <b>No</b> QA3.N-methylol derivatives <b>No</b> QA4.Monohaloalkene <b>No</b> QA5.S or N mustard <b>No</b> QA6.Propiolactones and propiosultones <b>No</b> QA7.Epoxides and aziridines <b>No</b> QA8.Aliphatic halogens <b>No</b>

Done

# Information Gathering & Evaluation Process

Chemical compounds - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://apps.ideaconsult.net:8180/ambit2/query/smarts?type=smiles&search=[\*]OC(=O)[%236%3BH1]%3D[%236%3BH1]c1cccc1&t

Disable Cookies CSS Forms Images Information Miscellaneous Outline Resize Tools View Source Options

Chemical compounds

ToxPredict TTC Depiction Datasets Chemical compounds Similarity Substructure Algorithms References Features Templates Models Ontology RDF playground Help

**ambit**

SMARTS

Keywords

Search for substructure and properties  
This site and AMBIT REST services are under development!

Retrieve data

Search results SMARTS [\*]OC(=O)[#6: Download as Max number of hits: 100

#	Compound	ECHA REGISTRATION DATE	ECHA CasRN	ECHA EC	ECHA Names	ECHA SYNON Names	ECHA SYNON Names	ECHA SYNON Names	ECHA SYNON Names	ECHA SYNON Names	ECHA SYNON Names
1		<a href="#">30.11.2010</a>	<a href="#">78-37-5</a>	<a href="#">201-110-3</a>	<a href="#">linalyl cinnamate</a>						
2		<a href="#">30.11.2010</a>	<a href="#">90-50-6</a>	<a href="#">201-999-8</a>	<a href="#">3,4,5-trimethoxycinnamic acid</a>						

http://apps.idea

# Information Gathering & Evaluation Process

[ToxPredict](#) [TTC Depiction](#) [Datasets](#) [Chemical compounds](#) [Similarity](#) [Substructure](#) [Algorithms](#) [References](#) [Features](#) [Templates](#) [Models](#) [Ontology](#) [RDF playground](#) [Help](#)

**ambit**


*This site and AMBIT REST services are under development!*

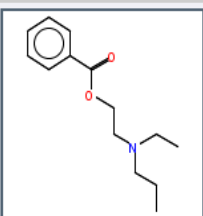
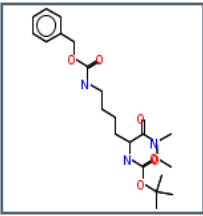
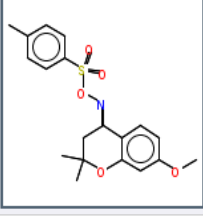
## Features

Find	Name	Units	Same as	Origin (Dataset, Model or Algorithm)	Nominal values
	<a href="#">Structural Alert for genotoxic carcinogenicity</a>		<a href="http://www.opentox.org/echaEndpoints.owl#Carcinogenicity">http://www.opentox.org/echaEndpoints.owl#Carcinogenicity</a>	<a href="#">Benigni / Bossa rulebase (for mutagenicity and carcinogenicity)</a>	NO
	<a href="#">Structural Alert for nongenotoxic carcinogenicity</a>		<a href="http://www.opentox.org/echaEndpoints.owl#Carcinogenicity">http://www.opentox.org/echaEndpoints.owl#Carcinogenicity</a>	<a href="#">Benigni / Bossa rulebase (for mutagenicity and carcinogenicity)</a>	NO
	<a href="#">Potential <i>S. typhimurium</i> TA100 mutagen based on QSAR</a>		<a href="http://www.opentox.org/echaEndpoints.owl#Carcinogenicity">http://www.opentox.org/echaEndpoints.owl#Carcinogenicity</a>	<a href="#">Benigni / Bossa rulebase (for mutagenicity and carcinogenicity)</a>	NO
	<a href="#">Unlikely to be a <i>S. typhimurium</i> TA100 mutagen based on QSAR</a>		<a href="http://www.opentox.org/echaEndpoints.owl#Carcinogenicity">http://www.opentox.org/echaEndpoints.owl#Carcinogenicity</a>	<a href="#">Benigni / Bossa rulebase (for mutagenicity and carcinogenicity)</a>	NO
	<a href="#">Potential carcinogen based on QSAR</a>		<a href="http://www.opentox.org/echaEndpoints.owl#Carcinogenicity">http://www.opentox.org/echaEndpoints.owl#Carcinogenicity</a>	<a href="#">Benigni / Bossa rulebase (for mutagenicity and carcinogenicity)</a>	NO
	<a href="#">Unlikely to be a carcinogen based on QSAR</a>		<a href="http://www.opentox.org/echaEndpoints.owl#Carcinogenicity">http://www.opentox.org/echaEndpoints.owl#Carcinogenicity</a>	<a href="#">Benigni / Bossa rulebase (for mutagenicity and carcinogenicity)</a>	NO
	<a href="#">For a better assessment a QSAR calculation could be applied.</a>		<a href="http://www.opentox.org/echaEndpoints.owl#Carcinogenicity">http://www.opentox.org/echaEndpoints.owl#Carcinogenicity</a>	<a href="#">Benigni / Bossa rulebase (for mutagenicity and carcinogenicity)</a>	NO
	<a href="#">Negative for genotoxic carcinogenicity</a>		<a href="http://www.opentox.org/echaEndpoints.owl#Carcinogenicity">http://www.opentox.org/echaEndpoints.owl#Carcinogenicity</a>	<a href="#">Benigni / Bossa rulebase (for mutagenicity and carcinogenicity)</a>	NO
	<a href="#">Negative for nongenotoxic carcinogenicity</a>		<a href="http://www.opentox.org/echaEndpoints.owl#Carcinogenicity">http://www.opentox.org/echaEndpoints.owl#Carcinogenicity</a>	<a href="#">Benigni / Bossa rulebase (for mutagenicity and carcinogenicity)</a>	NO
	<a href="#">Error when applying the decision tree</a>		<a href="http://www.opentox.org">http://www.opentox.org</a>	<a href="#">Benigni / Bossa rulebase (for mutagenicity and carcinogenicity)</a>	NO

<http://pirin.uni-plovdiv.bg:8080/malaria/feature/259>

# Information Gathering & Evaluation Process

Search results Property = Low (Class I) Download as  Max number of hits: 100

#	Compound	nature0910	nature0910	Benigni /	Benigni /
		Percentage inhibition 3D7	Percentage inhibition HEPG2	Structural Alert for genotoxic carcinogenicity	Structural Alert for nongenotoxic carcinogenicity
504		99.0	2.0	NO	NO
509		98.0	5.0	YES	NO
508		97.0	10.0	NO	NO
512		96.0	30.0	NO	NO

Done

# Information Requirements - *in vitro*

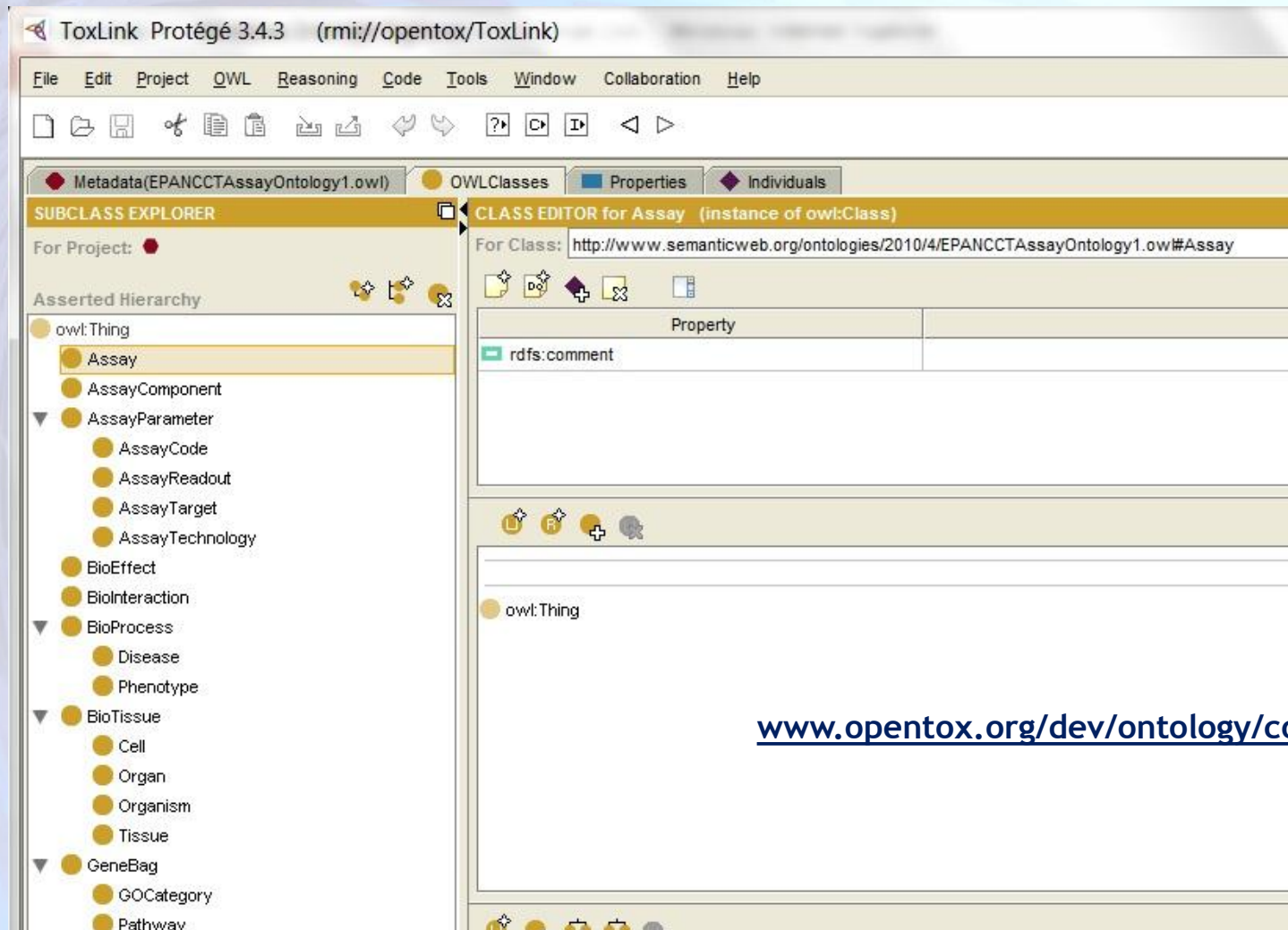
Two categories of *in vitro* methods are currently referred to within REACH as suitable:

- Validated methods. Examples include *in vitro* tests for skin corrosion and *in vitro* genotoxicity tests such as the Ames Salmonella typhimurium mutagenicity test.
- *In vitro* tests that meet internationally agreed pre-validation criteria from e.g. ECVAM.

A main criterion for acceptance for regulatory use is the adequacy of the information generated in such an *in vitro* assay for the purpose of classification and labeling and/or risk assessment.



# ToxLink: ToxCast Ontology

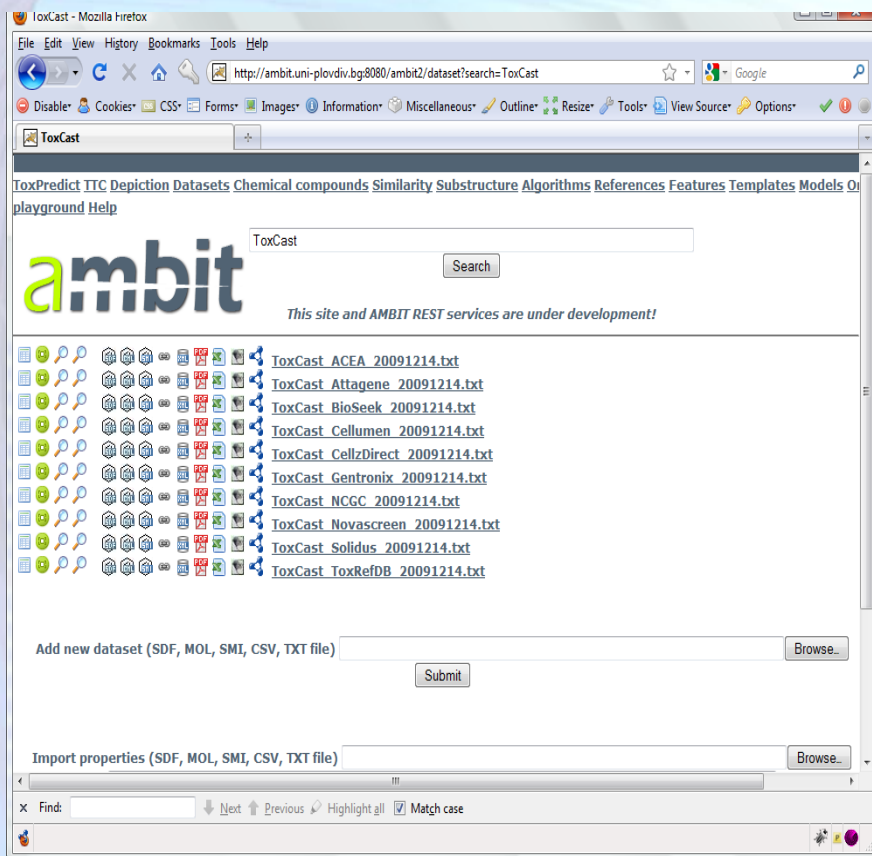


[www.opentox.org/dev/ontology/collaborative\\_protege](http://www.opentox.org/dev/ontology/collaborative_protege)

# Example: ToxCast

Dataset service at  
<http://ambit.uni-plovdiv.bg:8080/ambit2/dataset>

Query an OpenTox ontology service at  
<http://ambit.uni-plovdiv.bg:8082/ontology>



```
PREFIX ot:<http://www.opentox.org/api/1.1#>
PREFIX ota:<http://www.opentox.org/algorithms.owl#>
PREFIX owl:<http://www.w3.org/2002/07/owl#>
PREFIX dc:<http://purl.org/dc/elements/1.1/>
PREFIX rdfs:<http://www.w3.org/2000/01/rdf-schema#>
PREFIX rdf:<http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX otee:<http://www.opentox.org/echaEndpoints.owl#>
PREFIX toxcast:<http://www.opentox.org/toxcast#>
```

```
select *
where {
  ?Feature rdf:type ot:Feature.
  {?Feature dc:title ?title}.
  {?Feature owl:sameAs ?assay}.
  {?assay toxcast:gene ?geneid}.
  {?assay toxcast:hasProperty ?species}.
  {?species rdfs:type toxcast:SPECIES}.
  {?assay toxcast:hasProperty ?target_source}.
  {?target_source rdfs:type toxcast:ASSAY_TARGET_SOURCE}.
  {?assay toxcast:hasProperty ?target_family}.
  {?target_family rdfs:type toxcast:ASSAY_TARGET_FAMILY}.
  {?assay toxcast:hasProperty ?target}.
  {?target rdfs:type toxcast:ASSAY_TARGET}.
  {?assay toxcast:hasProperty toxcast:Cytochrome_P450}.
}
order by ?feature ?assay ?target
```

PREFIX ot:<<http://www.opentox.org/>  
 PREFIX ota:<<http://www.opentox.org/>  
 PREFIX owl:<<http://www.w3.org/2002>  
 PREFIX dc:<<http://purl.org/dc/element/>  
 PREFIX rdfs: <<http://www.w3.org/2000>  
 PREFIX rdf:<<http://www.w3.org/1999>  
 PREFIX otee:<<http://www.opentox.org/>  
 PREFIX toxcast:<<http://www.opentox.org/>

```

    select ?Feature ?title ?id
    where {
        ?Feature rdf:type ot:Feature
    }
    {?Feature dc:title ?title}.
    {?Feature owl:sameAs ?assay}.
    {?assay toxcast:gene ?geneid}.
    {?assay toxcast:hasProperty ?genename}.
    {?genename rdf:type toxcast:GENE_ID}
}

```

Query an OpenTox ontology service  
<http://ambit.uni-plovdiv.bg:8082>

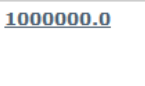

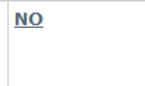

Chemical compounds

<http://ambit.uni-plovdiv.bg:8080/ambit2/datas>

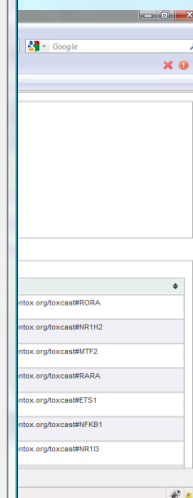
[Apple](#)
[Yahoo!](#)
[Google Maps](#)
[YouTube](#)
[Wikipedia](#)
[News \(114\)](#)
[Share on FriendFeed](#)

**Search results** Dataset = 961
 [Download as](#)
[Max number of](#)

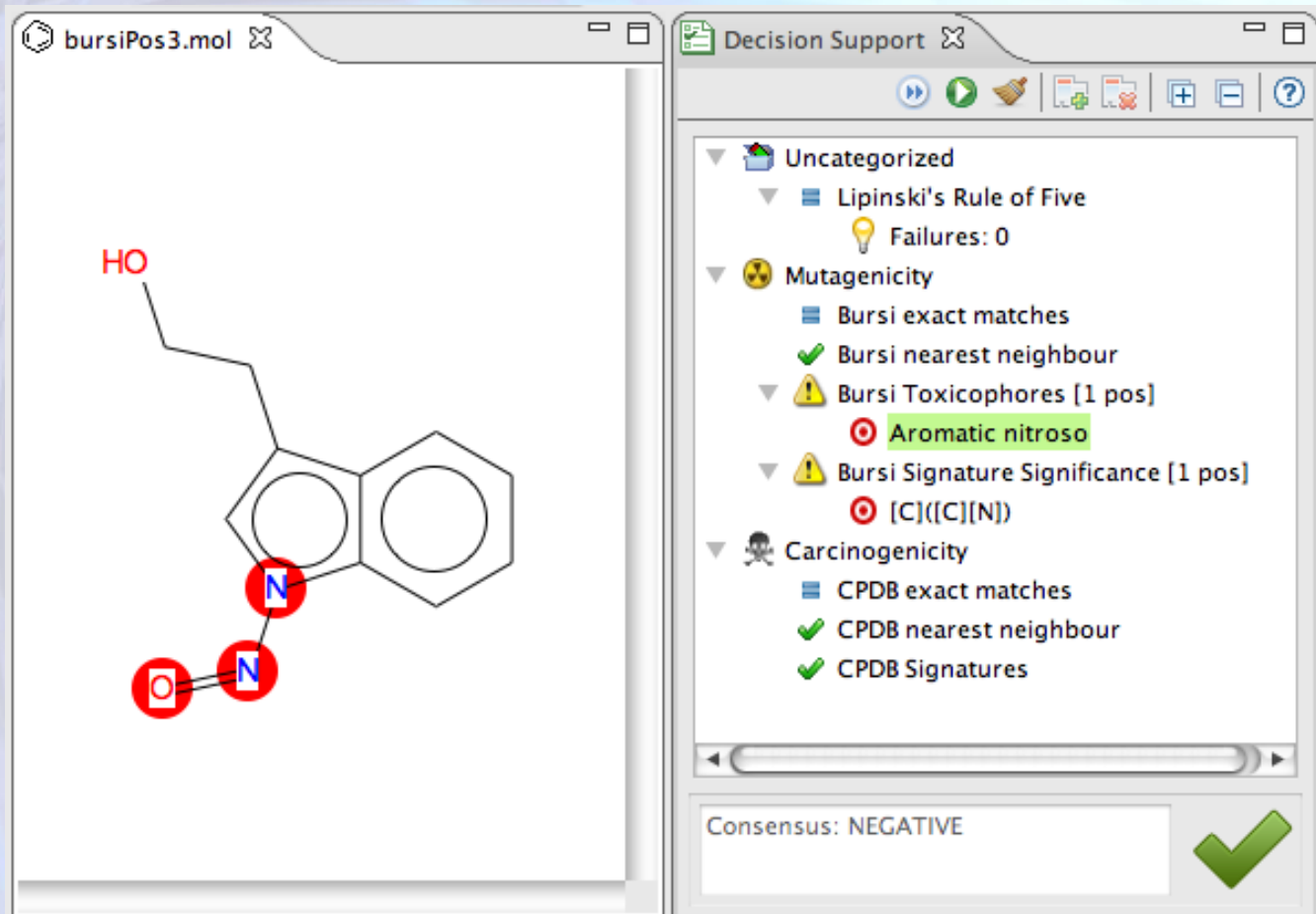
hits: 100

#	Compound	ToxCast At	Benigni /	Benigni /
		ATG RORE CIS	Structural Alert for genotoxic carcinogenicity	Structural Alert for nongenotoxic carcinogenicity
1		<a href="#">1000000.0</a>	<a href="#">NO</a>	<a href="#">NO</a>
2		<a href="#">1000000.0</a>	<a href="#">NO</a>	<a href="#">NO</a>
3		<a href="#">1000000.0</a>	<a href="#">NO</a>	<a href="#">NO</a>
4		<a href="#">1000000.0</a>	<a href="#">NO</a>	<a href="#">NO</a>

dataset/961?feat  
uni-  
feature/335126

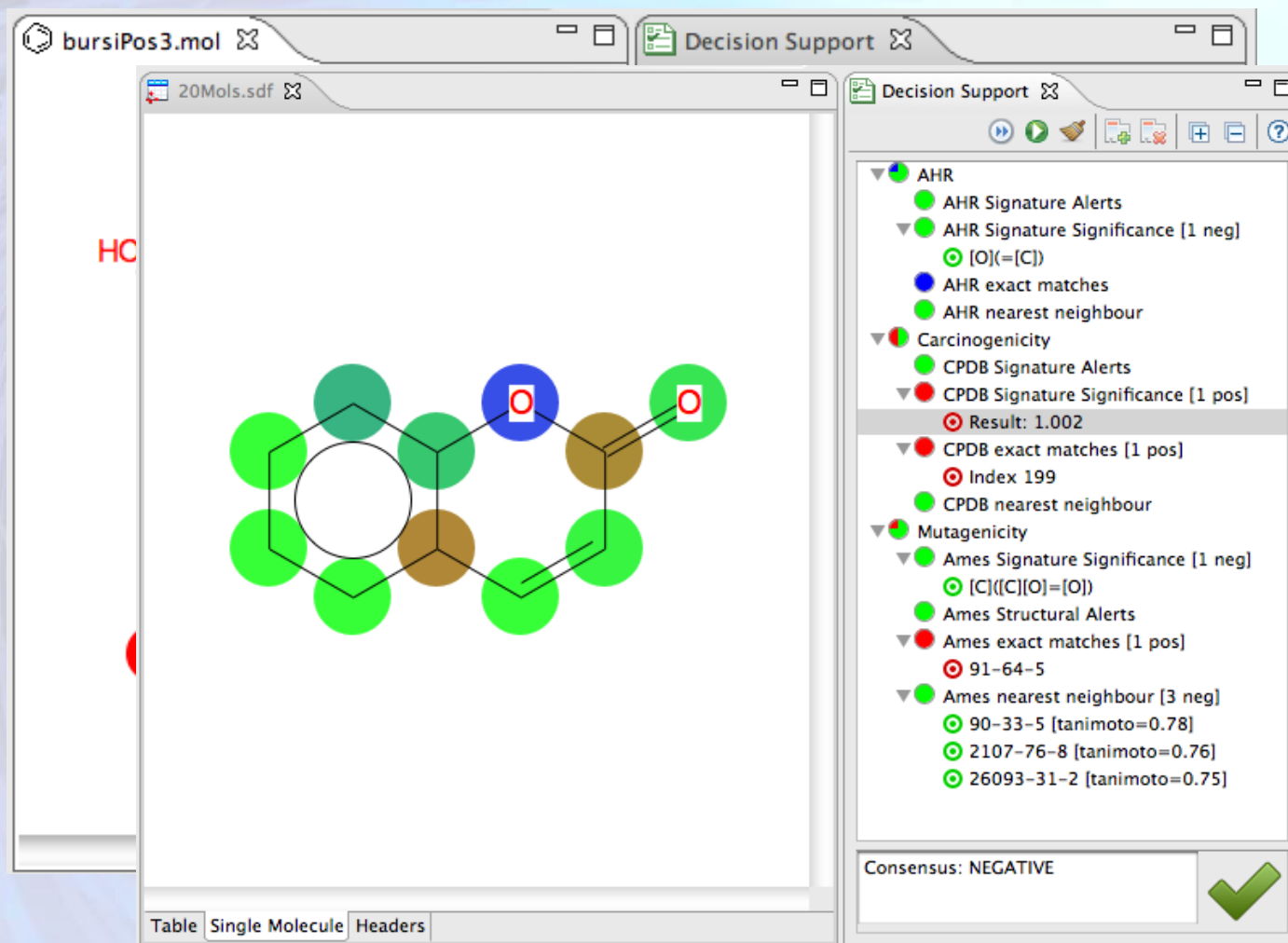


# Bioclipse Visualisation Workbench



O. Spjuth, L. Carlsson, M. Eklund, E. Ahlberg Helgee, and Scott Boyer.  
*Integrated decision support for assessing chemical liabilities. In preparation*

# Bioclipse Visualisation Workbench



O. Spjuth, L. Carlsson, M. Eklund, E. Ahlberg Helgee, and Scott Boyer.  
*Integrated decision support for assessing chemical liabilities. In preparation*



# Bioclipse Visualisation Workbench - OpenTox

HC

The screenshot displays the Bioclipse Visualisation Workbench interface. The main window shows a chemical structure of a steroid-like molecule with an epoxide group highlighted in red. The word "Changed" is written in purple above the structure. The interface includes several panels:

- Decision Support** (top right): A panel showing decision support results, including Ames Structural Alerts, Ames exact matches, Ames nearest neighbour, and OpenTox results.
- Properties** (bottom left): A table showing properties and values for the molecule.
- 2D-Structure** (bottom right): A panel showing the 2D structure of the molecule.

**Decision Support** results:

- Ames Structural Alerts [1 pos]
  - Epoxide
- Ames exact matches [no hits]
- Ames nearest neighbour [3 pos, 1 neg]
  - 26761-45-5 [tanimoto=0.82]
  - 2461-18-9 [tanimoto=0.81]
  - 2461-15-6 [tanimoto=0.73]
  - 5926-90-9 [tanimoto=0.71]
- OpenTox
  - Caco-2 Cell Permeability <http://www.n>  
caco2 = -4.548099994659424
  - Lipinski Rule of Five  
LipinskiFailures = 0.0
  - MolecularWeight

**Properties** table:

Property	Value
General	
Classification	POSITIVE
Matching atoms	22, 21, 23
Name	Epoxide
Test	Ames Structural Alerts

# Chemical Space Visualisation (Ches-Mapper)



## Ches-Mapper: Chemical Space Mapping and Visualization in 3D

<http://opentox.informatik.uni-freiburg.de/ches-mapper>



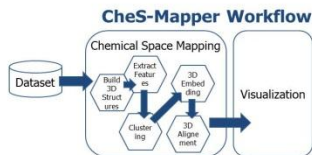
Martin Gütlein<sup>1\*</sup>, Andreas Karwath<sup>1</sup>, Stefan Kramer<sup>2</sup>

\*guetlein@informatik.uni-freiburg.de

<sup>1</sup>Institute for Computer Science • Albert-Ludwigs-Universität Freiburg • Germany, <sup>2</sup>Institute for Computer Science I12 • Technische Universität München • Germany

### Abstract

Scientific researchers in the field of cheminformatics, are often overwhelmed by the size and the sheer complexity of chemical datasets. Therefore, the need for visualization tools, is one of the utmost requests. Our recently developed 3D molecular viewer Ches-Mapper (Chemical Space Mapper) includes many techniques, like state-of-the-art structural clustering, and multi-dimensional embedding techniques. Large datasets are divided into clusters of similar compounds and consequently arranged in 3D space, such that their spatial proximity reflects their chemical similarity. This intuitively provides essential information to the user, while making the dataset more easily accessible and allowing easy and understandable access to a large number of chemical structures within seconds. The different clustering approaches employed in our tool utilize common substructures as well as quantitative chemical descriptors of the compounds. These features can be highlighted within Ches-Mapper, which aids the chemist to better understand the underlying scientific knowledge. As a final function, the tools can also be used to select and export specific part of a given dataset for further analysis.



### Wizard Dialog to Control Mapping

- A wizard dialog guides through the Mapping process
- Suitable for novice and expert users
- Single Steps:
  - Load dataset
  - Build 3D structure
  - Extract features
  - Clustering
  - 3D Embedding
  - 3D Alignment
- Automatic detection and plug in of new methods and algorithms



### Chemical Space Mapping

#### Build 3D Structure and Extract Features

- Select input dataset
  - Various dataset formats are supported (csv/xml/indirect...)
  - Dataset can be directly loaded from the web
- 3D structure is built
  - 3D structure can be built with Chemical Development Kit (CDK) or OpenBabel
  - External libraries like Corina can be plugged in easily
- Extract features
  - Features are required for clustering and embedding
  - Automatic extraction of dozens of descriptors with CDK



Feature	Value
LogP	8.4281
MolWeight	281.408
MolWeight	281.4
Molecular weight	281.4

#### 3D Embedding (of Clusters & Compounds)

- Embedding algorithms assign 3D coordinates to each compound or cluster, according to the feature values of the compounds
- Different approaches are provided:
  - Principal Component Analysis (PCA)
  - Multi-dimensional Scaling Using Majorization (SMACOF)
  - T-distributed Stochastic Neighbor Embedding (tSNE)
- Developers can easily plug in their own preferred 3D Embedding algorithm

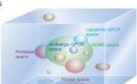
Features	LogP	Weight	Wt path number	Wt priority number
CCCCCCCCCCCCCCCCCCCC	4.27	144.064	144.0	27.0
CCCCCCCCCCCCCCCCCCCC	2.74	113.962	113.0	13.0
CCCCCCCCCCCCCCCCCCCC	4.27	144.064	144.0	27.0

NO	X	Y	Z
CCCCCCCCCCCCCCCCCCCC	-20.48	16.99	30.45
CCCCCCCCCCCCCCCCCCCC	16.99	30.45	16.99
CCCCCCCCCCCCCCCCCCCC	2.94	-16.99	-30.45

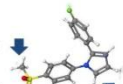
#### Cluster Compounds

- Compounds in the dataset are assigned to subgroups according to their similarity
- Supported cluster algorithms:
  - k-Means Clustering
    - Fixed number of k clusters
    - Random initialization, iterative update of clusters and cluster centroids
  - Hierarchical Clustering
    - Each compound is single cluster
    - Sequentially merge similar clusters
  - Structural Clustering
    - Find groups that share structural similarity
    - Compounds are assigned to clusters when there exists a common subgraph of sufficient size
- Developers can plug in new cluster algorithms



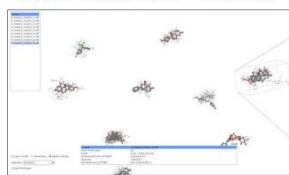
#### 3D Alignment of Compounds

- Compounds in a cluster are likely to share common subgraphs:
  - This subgraph is already available if structural clustering is performed
  - Alternatively, the maximum common subgraph can be computed within each cluster
- The compounds within a cluster can be superimposed/aligned according to this subgraph:
  - This shows differences between compounds



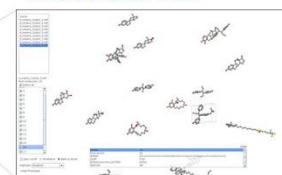
### 3D Visualization

#### Dataset Overview -- Clusters



- Datasets are separated into clusters, arranged in 3D space
- The intuitive interface of the 3D viewer allows to:
  - Zoom/rotate the clusters
  - Get valuable information on clusters via mouse over
  - Examine a cluster by clicking on it
- The embedding into 3D space (positions/distance between clusters) reflects the similarity between clusters
- Cluster can be removed from the dataset

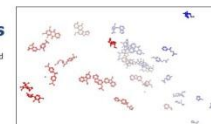
#### Inside Cluster View



- By selecting a cluster, the view zooms into the cluster and displays only the compounds included
  - Details for each compound are available via mouse over
  - Like the clusters, the compounds are embedded into the 3D space as well: the position/distance between compounds within the cluster reflects the similarity between compounds
- Compounds can be removed from the dataset

#### Highlight Features and Endpoints

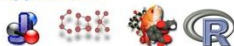
- All compound properties can be highlighted:
  - The compounds are colored according to the numeric value, a high value is indicated by red, a low value is indicated by blue
  - Also available for the cluster overview
- Gives an intuitive explanation towards the quality of the clustering approach:
  - Does the clustering algorithm separate active from inactive compounds?



### Open-Source Webstart Application

- Java program that comes in two variants:
  - Java Web Start application (can directly started from a web browser)
  - Local installation that makes use of non-Java libraries
- Ches-Mapper is available at <http://opentox.informatik.uni-freiburg.de/ches-mapper>

Powered by:



### References

- [1] Seifried, G., Griesch, T., Buchwald, F., Kramer, S. Online Structural Graph Clustering Using Frequent Subgraph Mining, 2010, Machine Learning and Knowledge Discovery in Databases, 213–228, Springer
- [2] Email an open-source Java viewer for chemical structures in 3D. <http://www.jmol.org>

### Acknowledgements

This work has been supported by the EU FP7 project (HEALTH-F5-2008-200787) OpenTox (<http://www.opentox.org>).

Developed by Martin Gütlein, Andreas Karwath, Stefan Kramer (ALU & TUM)

# Chemical Space Visualisation (Ches-Mapper)



## CheS-Mapper: Chemical Space Mapping and Visualization in 3D



### Abstract

Scientific researchers in the field of cheminformatics, chemical datasets. Therefore, the need for visualization. Our recently developed 3D molecular viewer Ches-Mapper of 3D structural clustering, and multi-dimensional or similar compounds and consequently arranged in 3D similarity. This intuitively provides essential information allowing easy and understandable access to a large number of different clustering approaches employed in our tool to descriptors of the compounds. These features can be understood the underlying scientific knowledge. As a first part of a given dataset for further analysis.

### Build 3D Structure and Extract Features

- Select input dataset
  - Various dataset formats are supported (csv/mvc/indirect...)
  - Dataset can be directly loaded from the web
- 3D structure is built
  - 3D structure can be built with Chemical Development Kit (CDK) or OpenBabel
  - External libraries like Corina can be plugged in easily
- Extract features
  - Features are required for clustering and embedding
  - Automatic extraction of dozens of descriptors with CDK

### Cluster Compounds

- Compounds in the dataset are assigned to subgroups according to their similarity
- Supported cluster algorithms:
  - k-Means Clustering
    - Fixed number of k clusters
    - Random initialization, iterative update of clusters and cluster centroids
  - Hierarchical Clustering
    - Each compound is single cluster
    - Sequentially merge similar clusters
  - Structural Clustering
    - Finds groups that share structural similarity
    - Compounds are assigned to clusters where there exists a common subgraph of sufficient size
- Developers can plug in new cluster algorithms

Home  
k\_means\_cluster\_1.sdf  
k\_means\_cluster\_2.sdf  
k\_means\_cluster\_3.sdf  
k\_means\_cluster\_4.sdf  
k\_means\_cluster\_5.sdf  
k\_means\_cluster\_6.sdf  
k\_means\_cluster\_7.sdf  
k\_means\_cluster\_8.sdf  
k\_means\_cluster\_9.sdf

☐ Spin on/off ☐ Wireframe ☒ Balls & Sticks

Highlight:

Initial Prototype

Cluster	k_means_cluster_9.sdf
Num molecules	19
LOGP	[2,81; 9,09] 0:5,09
ActivityOutcome_NCTRE	[ active(19) ]
Species	[ rat(19) ]
ActivityScore_NCTRE	[80; 100] 0:88,11

# Forming chemical feature-based categories



## CheS-Mapper: Chemical Space Mapping and Visualization in 3D



### Abstract

Scientific researchers in the field of chemoinformatics chemical datasets. Therefore, the need for visual: Our recently developed 3D molecular viewer CheS of 3D-structure clustering, and multi-dimens similar compounds and consequently arranged similarity. This intuitively provides essential info: allowing easy and understandable access to a larg The different clustering approaches employed in descriptors of the compounds. These features ca understand the underlying scientific knowledge, a part of a given dataset for further analysis.

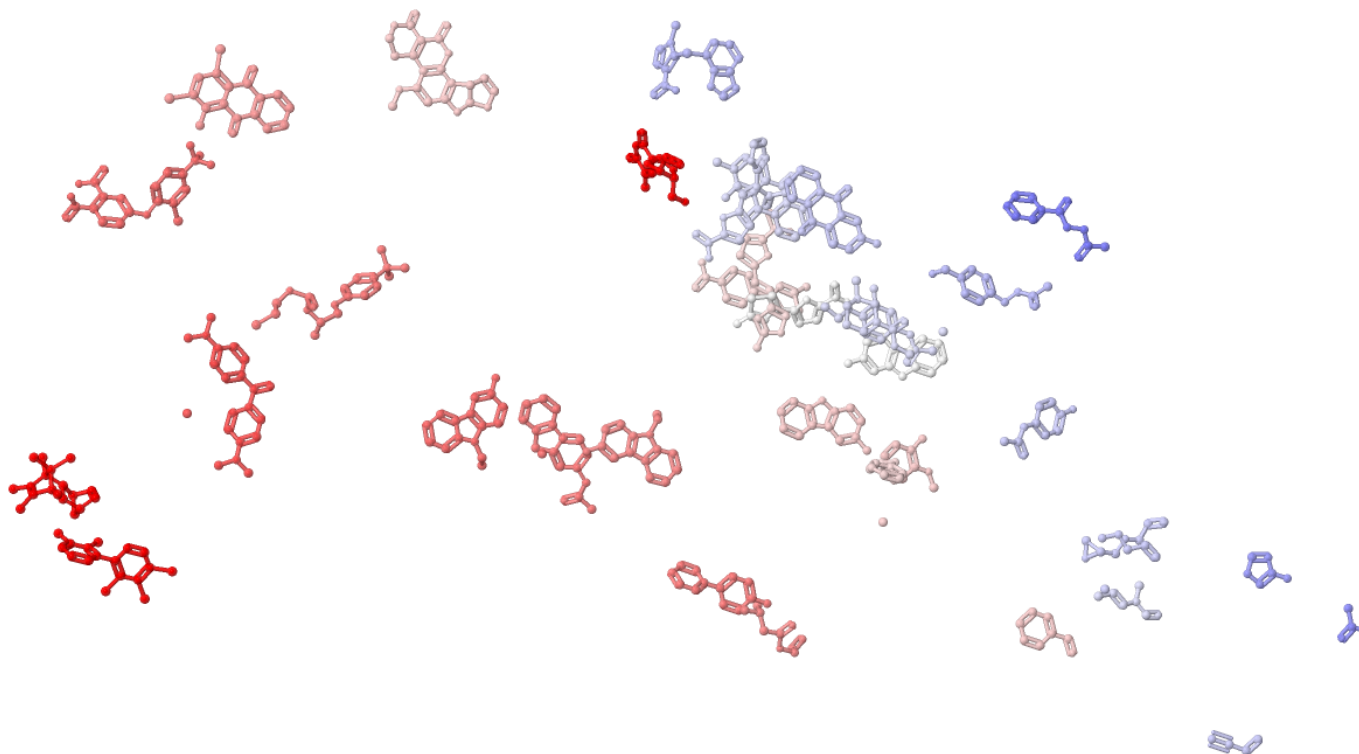
### Build 3D Structure Extract Features

- Select input dataset
  - Various dataset formats are supported (csv/xml/indirect...)
  - Dataset can be directly loaded from the web
- 3D structure is built
  - 3D structure can be built with Chemical Development-Kit (CDK) or OpenBabel
  - External libraries like Corina can be plugged in easily
- Extract features
  - Features are required for clustering and embedding
  - Automatic extraction of dozens of descriptors with CDK

### Cluster Compound

- Compounds in the dataset are assigned to their similarity
- Supported cluster algorithms:
  - k-Means Clustering
    - Fixed number of k clusters
    - Random initialization, iterative up clusters and cluster centroids
  - Hierarchical Clustering
    - Each compound is single cluster
    - Sequentially merge similar cluster
  - Structural Clustering
    - Finds groups that share structural
    - Compounds are assigned to cluster there exists a common subgraph
- Developers can plug in new cluster algorithm

- Java program that comes in two variants:
  - Java Web Start application (can direct
  - Local installation that makes use of no
- CheS-Mapper is available at <http://open.tox>



# Read-Across Application

[File](#) [Edit](#) [View](#) [History](#) [Bookmarks](#) [Tools](#) [Help](#)

[http://localhost:8085/vtox#Categories](#)

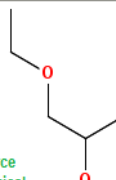
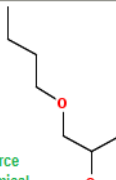
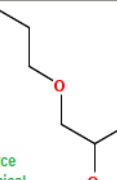
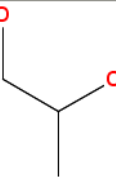
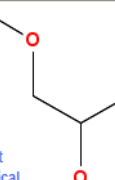
[Disable](#) [Cookies](#) [CSS](#) [Forms](#) [Images](#) [Information](#) [Miscellaneous](#) [Outline](#) [Resize](#) [Tools](#) [View Source](#) [Options](#)

[http://localhost:8085/vtox#Login](#)

## ToxPredict

[Help](#)

Page: 0 ☒ Show identifiers ☒ Show categories [Assign categories](#) [Download](#)

Property	1	2	3	4	5
 Source chemical	 Source chemical	 Source chemical		 Target chemical	
Select Structures	Yes	Yes	Yes	No	Yes
CASRN	1569-02-4	5131-66-8	1569-01-3	57-55-6	107-98-2
ChemicalName	1-ethoxypropan-2-ol	1-butoxypropan-2-ol	1-propoxypropan-2-ol	Propylene glycol	1-methoxypropan-2-ol
EINECS	216-374-5	225-878-4	216-372-4		203-539-1
IUPACName					
Categories					
Neutral Organics (alcohol)	YES	YES	YES	YES	YES
Neutral Organics (ether)	YES	YES	YES		YES

[Categories](#)

[WELCOME, GUEST](#)  
[My account](#)  
[Log out](#)  
[PREDICT](#)  
[Search structure](#)  
[Upload structure](#)  
[View results](#)  
[BROWSE](#)  
[Datasets](#)  
[Models](#)  
[READ ACROSS](#)  
[Target compound](#)  
[Analogues](#)  
[Categories](#)  
[Phys Chem properties](#)  
[Toxicity](#)  
[Read Across](#)  
[MY WORKSPACE](#)  
[My uploads](#)

[Twitter](#) [LinkedIn](#) [Facebook](#) [More...](#)

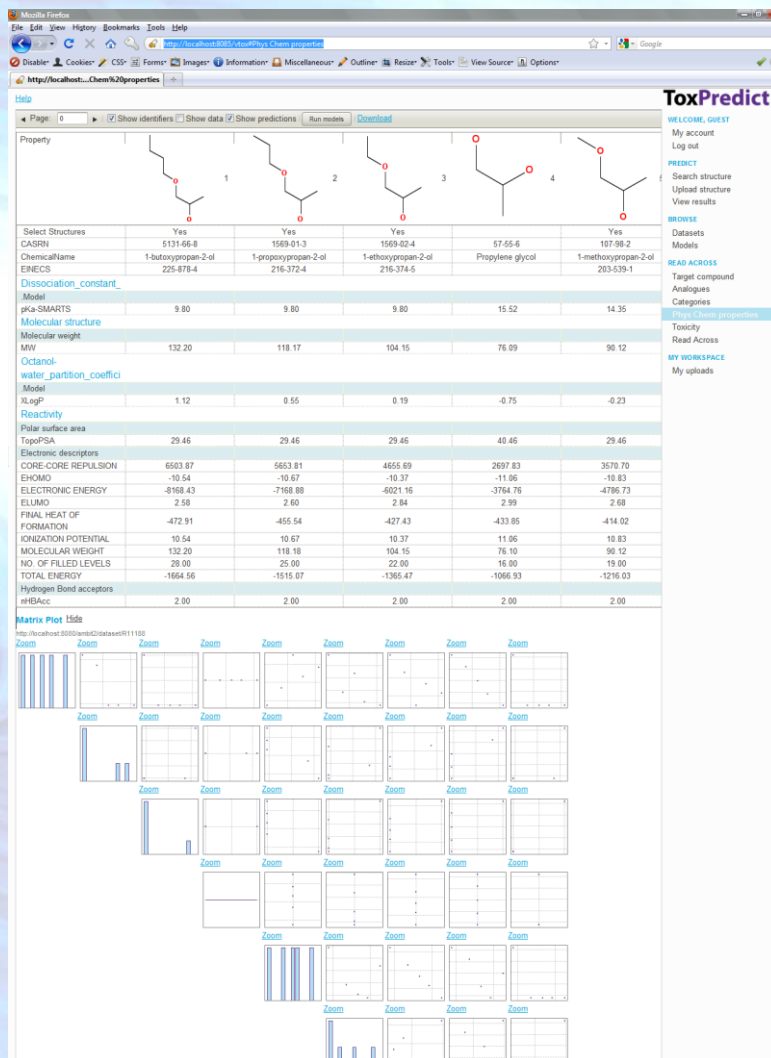
Developed by [Ideaconsult Ltd.](#) 2011

Find: skin [Next](#) [Previous](#) [Highlight all](#) [Match case](#)

Done



# Read-Across Application



# Read-Across Application

Mozilla Firefox  
http://localhost:8085/vtox#Toxicity

Disable Cookies CSS Forms Images Information Miscellaneous Outline Resize Tools View Source Options

http://localhost:8085/vtox#Login

Help

Page: 0 Show identifiers Show data Show predictions Run models Download

Property

1 2 3 4

**Carcinogenicity**

Dataset	1	2	3	4	5	6
CAS					57-55-6	
Canc						
Mouse_Female_Canc					NO	
Mouse_Female_NTP					NO	
Mouse_Male_Canc					NO	
Mouse_Male_NTP					NO	
Rat_Female_Canc						
Rat_Female_NTP					NO	
Rat_Male_Canc						
Rat_Male_NTP					NO	
Reference					NO	CPDB
SAL						
Model						
Potential S. typhimurium TA100 mutagen based on QSAR	NO	NO	NO	NO	NO	NO
Potential carcinogen based on QSAR	NO	NO	NO	NO	NO	NO
Structural Alert for genotoxic carcinogenicity	NO	NO	NO	NO	NO	NO
Structural Alert for nongenotoxic carcinogenicity	NO	NO	NO	NO	NO	NO
Unlikely to be a S. typhimurium TA100 mutagen based on QSAR	NO	NO	NO	NO	NO	NO
Unlikely to be a carcinogen based on QSAR	NO	NO	NO	NO	NO	NO

EPA Integrated Risk Information System (IRIS) Toxicity Review Data

Inhalation_RIC_Assessed					1.00
Inhalation_RIC_Confidence					Medium
Inhalation_RIC_CriticalEffects					mild reversible sedation
Inhalation_RIC_Notes					NOEL (No observed adverse effect level) HEC (Huma
Inhalation_RIC_mg_per_m3					2.00
Inhalation_RIC_mmol_per_m3					0.02
Inhalation_StudyRoute					
Inhalation_UnitRisk_Assesse					0.00
Inhalation_UnitRisk_microg_m3					
Inhalation_UnitRisk_micromc					

Physico chemical properties >> Boiling point

Dataset	1	2	3	4	5	6
BP					119.00	
Boiling Point Pressure						
CAS					107-98-2	
Error					-2.45	
EstBP					116.55	
ExpBP					119.00	
NAME					1-METHOXY-2-PROPANOL	

Skin sensitisation

Model	1	2	3	4	5	6
Alert for Acyl Transfer agent identified.	NO	NO	NO	NO	NO	NO
Alert for Michael Acceptor identified.	NO	NO	NO	NO	NO	NO
Alert for SN2 identified.	NO	NO	NO	NO	NO	NO
Alert for SN4 identified.	NO	NO	NO	NO	NO	NO
Alert for Schiff base formation identified.	NO	NO	NO	NO	NO	NO
No skin sensitisation alerts identified.	YES	YES	YES	YES	YES	YES

Matrix Plot Show

http://localhost:8085/ambid2/dataset/R1118

**ToxPredict**

WELCOME, GUE ST  
My account  
Log out

**PREDICT**  
Search structure  
Upload structure  
View results

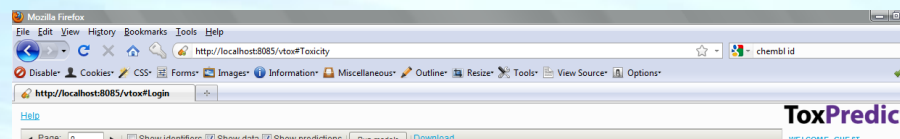
**BROWSE**  
Datasets  
Models

**READ ACROSS**  
Target compound  
Analogues  
Categories  
Phys Chem properties

**Toxicity**  
Read Across

**MY WORKSPACE**  
My uploads

# Read-Across Application



Carcinogenicity					
Dataset					
CAS				57-55-6	
Canc				YES	
Mouse_Female_Canc				ND	
Mouse_Female_NTP				ND	
Mouse_Male_Canc				ND	
Mouse_Male_NTP				ND	
Rat_Female_Canc				YES	
Rat_Female_NTP				ND	
Rat_Male_Canc				YES	
Rat_Male_NTP				ND	
Reference				CPDB	
SAL				YES	
Model					
Potential S. typhimurium TA100 mutagen based on QSAR	NO	NO	NO	NO	NO
Potential carcinogen based on QSAR	NO	NO	NO	NO	NO
Structural Alert for genotoxic carcinogenicity	NO	NO	NO	NO	NO
Structural Alert for nongenotoxic carcinogenicity	NO	NO	NO	NO	NO
Unlikely to be a S. typhimurium TA100 mutagen based on QSAR	NO	NO	NO	NO	NO
Unlikely to be a carcinogen based on QSAR	NO	NO	NO	NO	NO
EPA Integrated Risk Information System (IRIS) Toxicity Review Data					
Inhalation_RfC_Assessed					1.00
Inhalation_RfC_Confidence					Medium
Inhalation_RfC_CriticalEffects					mild reversible sedation
Identified	NO	NO	NO	NO	NO
Alert for SN2 identified	NO	NO	NO	NO	NO
Alert for SN4 identified	NO	NO	NO	NO	NO
Alert for Schiff base formation identified	NO	NO	NO	NO	NO
No skin sensitisation alerts identified	YES	YES	YES	YES	YES

# Ontology - automating consistency

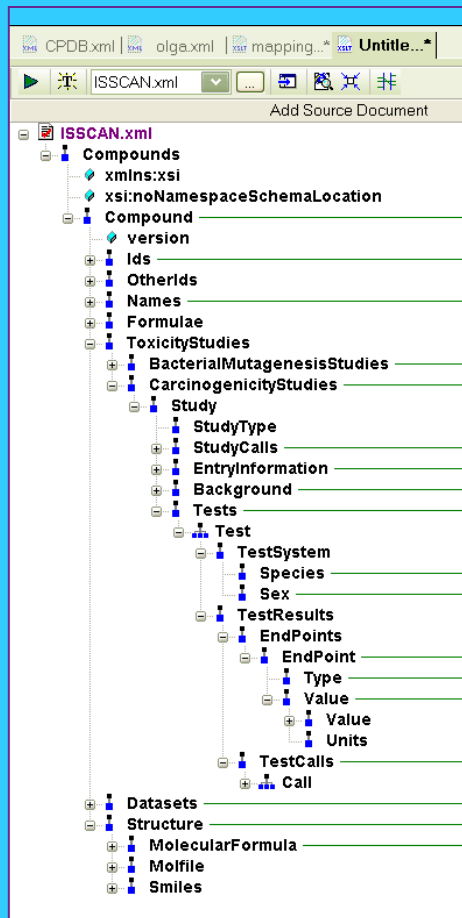
We have defined an ontology based on the OWL (Web Ontology Language) for toxicological endpoints which is in line with current ECHA REACH guidance.

Using this ontology, each attribute in a toxicological dataset can be associated with an entry to the ontology, therefore allowing a unique mapping between endpoints in various and heterogeneous datasets.

The mapping of chemical compound properties, stored in the OpenTox prototype database, with the endpoints ontology, and the information which properties are predicted by models, available via the OpenTox model service, is used to automatically recognise which endpoints have predictive models available, and ensures consistency of the used endpoint terminology across the set of distributed OpenTox services.

# Toxicological Endpoint Ontology Development

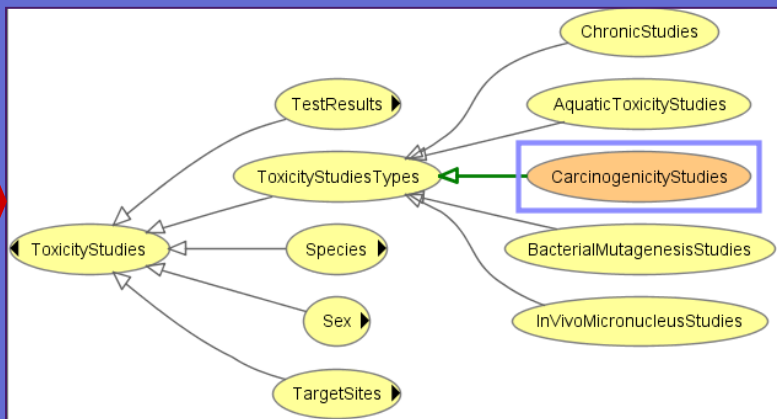
## ToxML schema



Other publicly available resources:  
DSSTox, GoReni (ITEM), ISSCAN ...

OpenTox  
Toxicological  
Endpoint  
Ontology

## Ontology Development



Re-use of terms defined in  
neighbouring ontologies (e.g. OBO)

Collaborative  
Protégé  
Environment



# OpenToxipedia



Barry Hardy Log out Quicktools Site Setup Help

Site Map Accessibility Contact Data

Search Site

Home Toxicity Prediction OpenTox Blog People Partners Development OpenToxipedia

User Guidance Latest Entries A B C D E F G H I J K L M N O P Q R S T U V W  
X Y Z by Categories Entries OpenToxipedia

You are here: Home » OpenToxipedia

Contents View Edit Rules Sharing History

Actions Display Add new... State: Published

## OpenToxipedia

by Barry Hardy — last modified Sep 03, 2009 01:09 PM

### OpenTox Community Resource for Toxicology Vocabulary and Ontology

OpenTox is supporting the creation and curation of OpenToxipedia, a community-based predictive toxicology knowledge resource. All members of the community are welcome to provide entries, suggested definition edits or additional information to entries in the resource.

OpenTox is supporting the application and development of the **ToxML** standard for representation of toxicology data, the **OECD principles for (Q)SAR model validation**, and the use of the **OECD HT** standard for regulatory reporting purposes.

OpenToxipedia provides here a Vocabulary Resource of toxicology terminology. We hope you find the resource useful and consider contributing to terms and their content.

### Guidance for Vocabulary Resource entries



[www.opentox.org/opentoxipedia](http://www.opentox.org/opentoxipedia)

# A Toxicology Ontology Roadmap

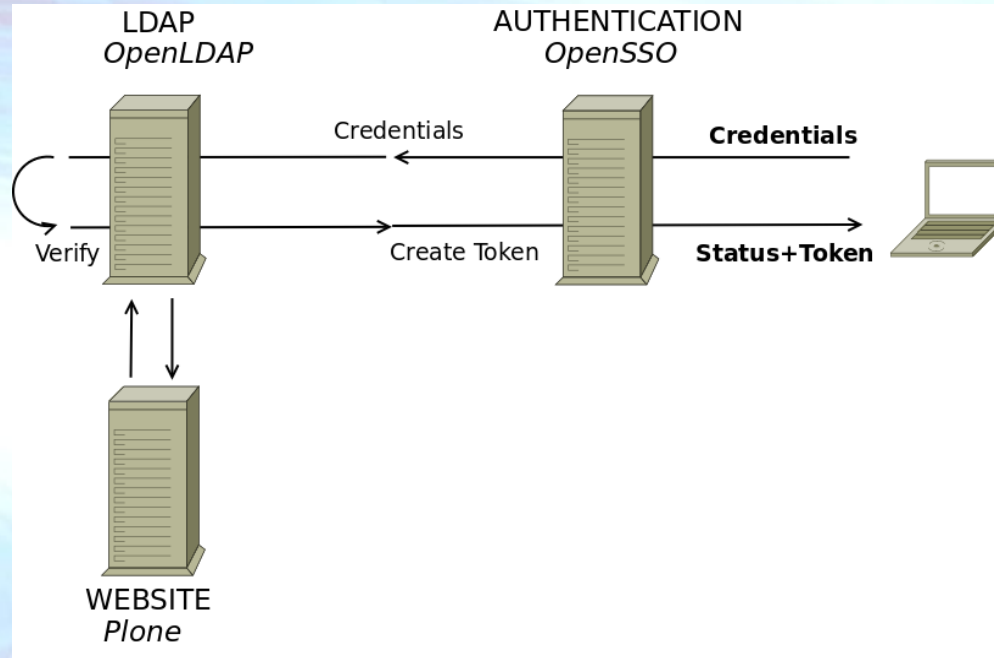
*In Preparation by:* Barry Hardy (Douglas Connect and OpenTox), Ian Dix (AstraZeneca & Pistoia Alliance), Sherri Matis-Mitchell (AstraZeneca), David Cook (AstraZeneca), David Heard (Novartis), Dominic Clark (EMBL-EBI), John Overington (EMBL-EBI), Philip Judson (Lhasa), David Watson (Lhasa), Anne Hersey (EMBL-EBI), Andrew White (Unilever), Loca Toldo (Merck KGaA), Gordana Apic (Cambridge Cell Networks), Imran Shah (US EPA), Chihae Yang (Altamira), Dave Bower (Leadscope), Ola Spjuth (Univ Uppsala), Janna Hastings (EMBL-EBI), Philip Carthew (Unilever), ----

Based on Proceedings from the Toxicology Ontology Roadmap Workshop  
EMBL-EBI Industry Programme Workshop  
16 -17th November 2010, Hinxton, UK

# Controlling Access to Confidential Information

- OpenTox makes resources available through **URIs**
- OpenTox provides facilities to protect confidential information located at **URIs**. Two tasks are involved here:
  - **Authentication**: Confirming the identity of the user requesting access
  - **Authorisation**: Granting the confirmed identity access according to a set of restrictions described in policies

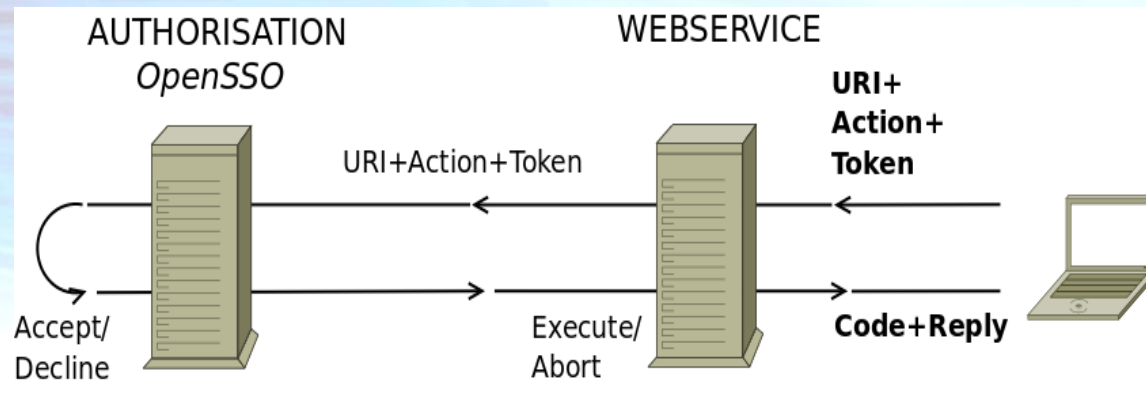
# Authentication



- Registered users are instantly available as potential users of OpenTox web services
- Users receive a token upon service request



# Authorisation



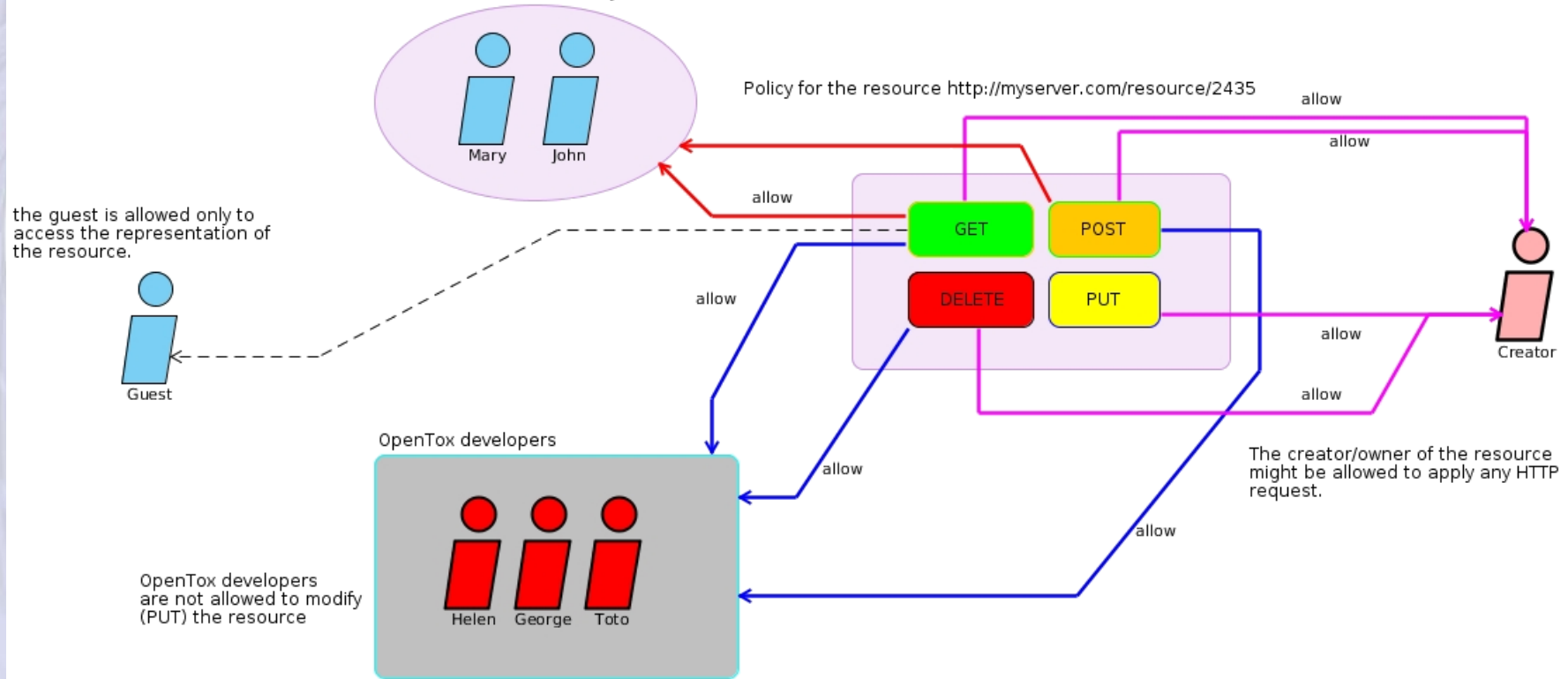
- Tokens encode user identity
- Tokens are valid for a certain time period only (customizable)
- The triplet URI+Action+Token makes up the call to be authorised
- All messages are encrypted (SSL)
- Resource Owners create and modify policies defining access rules



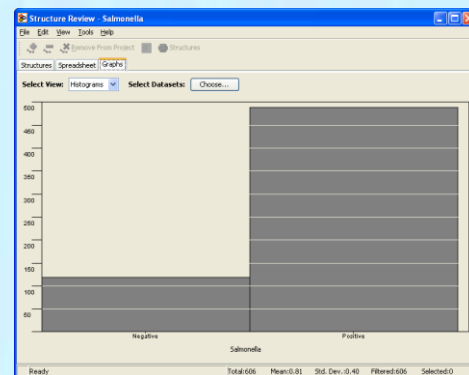
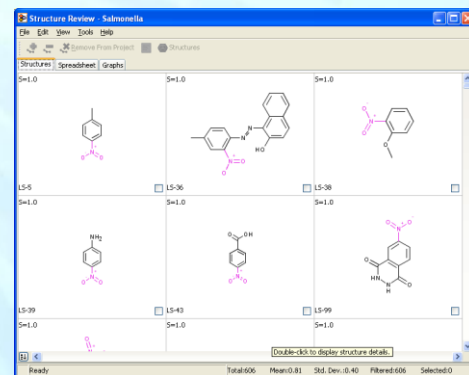
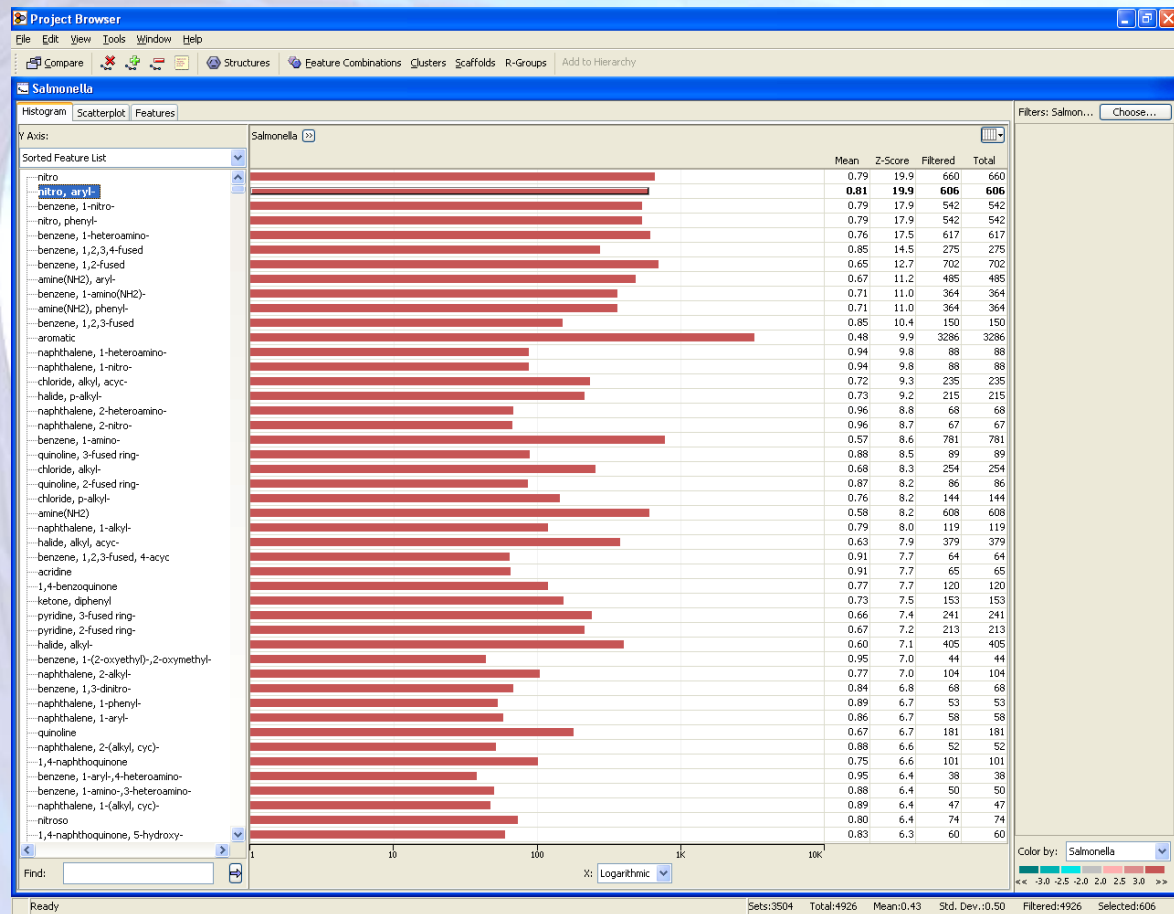
# Policies

Visual Paradigm for UML Community Edition [not for commercial use]

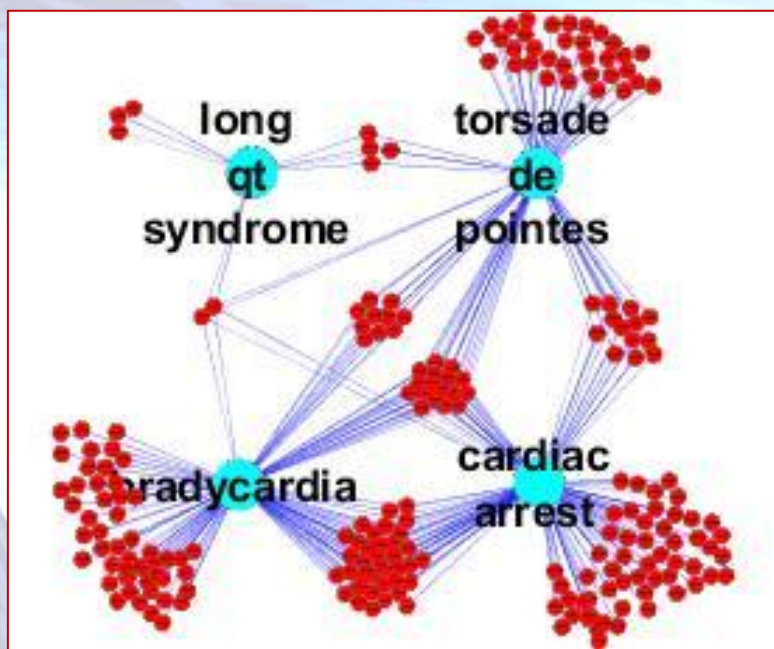
Chemistry Lab Staff



# OpenTox - Leadscope Integration



# Analysis of Adverse Events Based on Pharmacological Activity



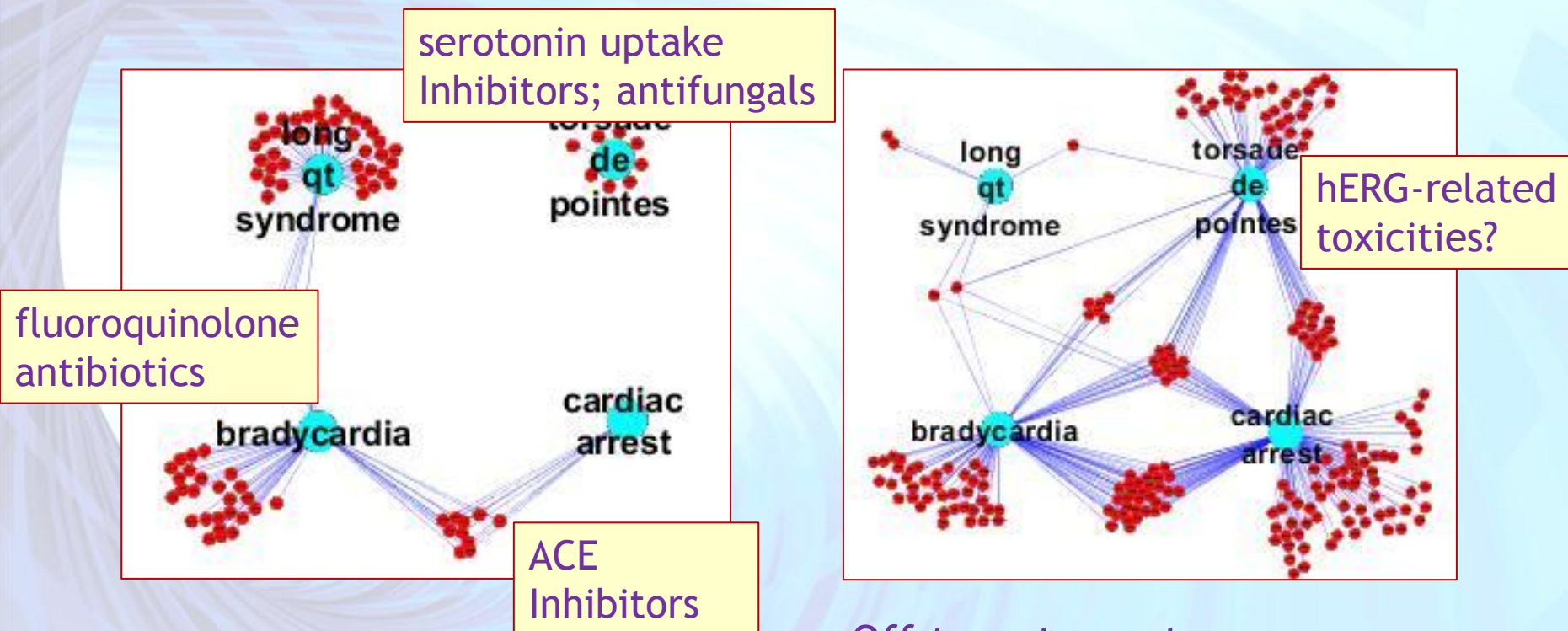
- Question addressed:
  - Are the adverse events a function of inhibiting the pharmacological target?
  - Or is the adverse event due to an off-target activity?



- Cardiac adverse events
- Related to hERG ion channel?

cyan = adverse event, red = drug  
lines define links

# Example: Cardiac Adverse Events



On-target events

cyan = adverse event, red = drug  
lines define links

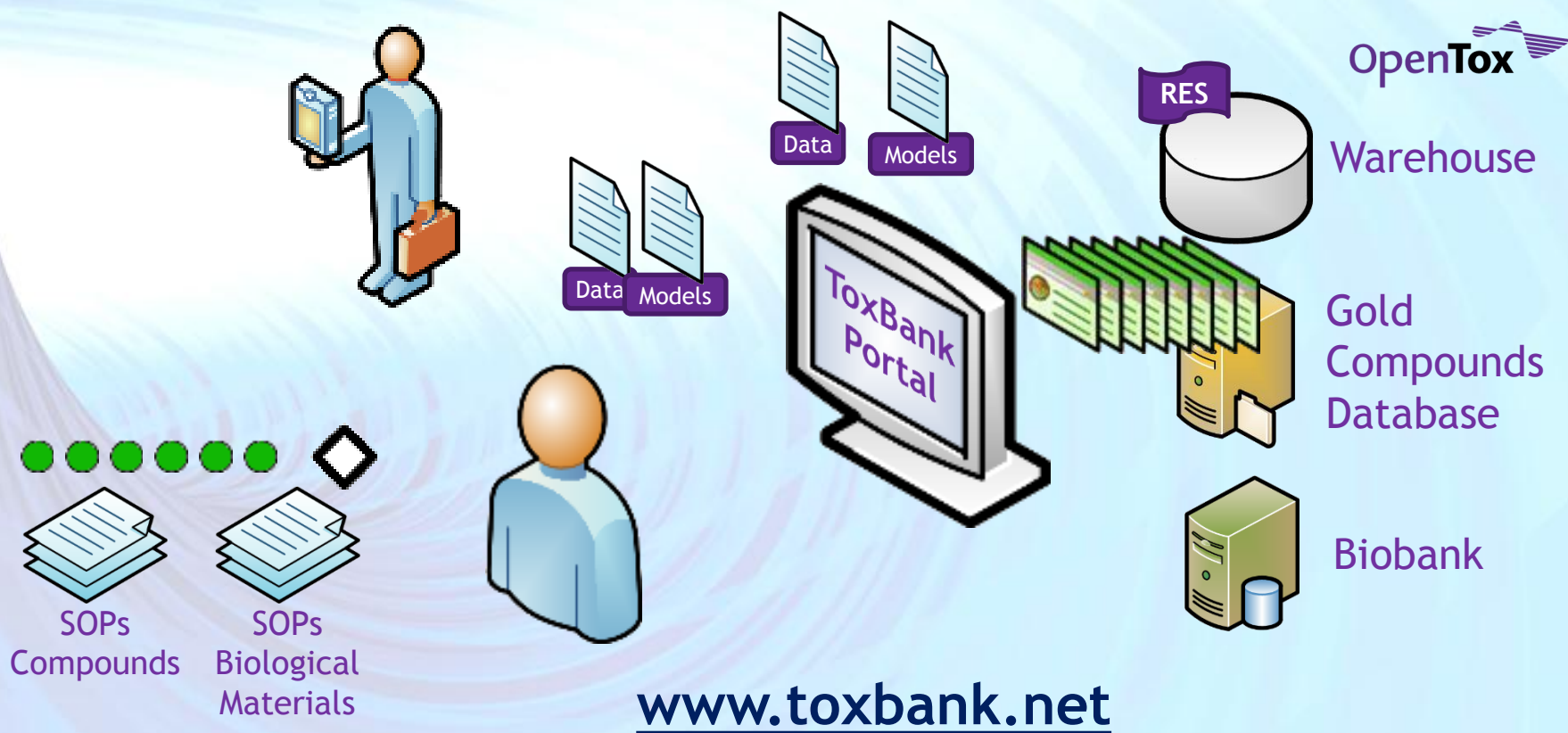
Off-target events

PHARMATROPE



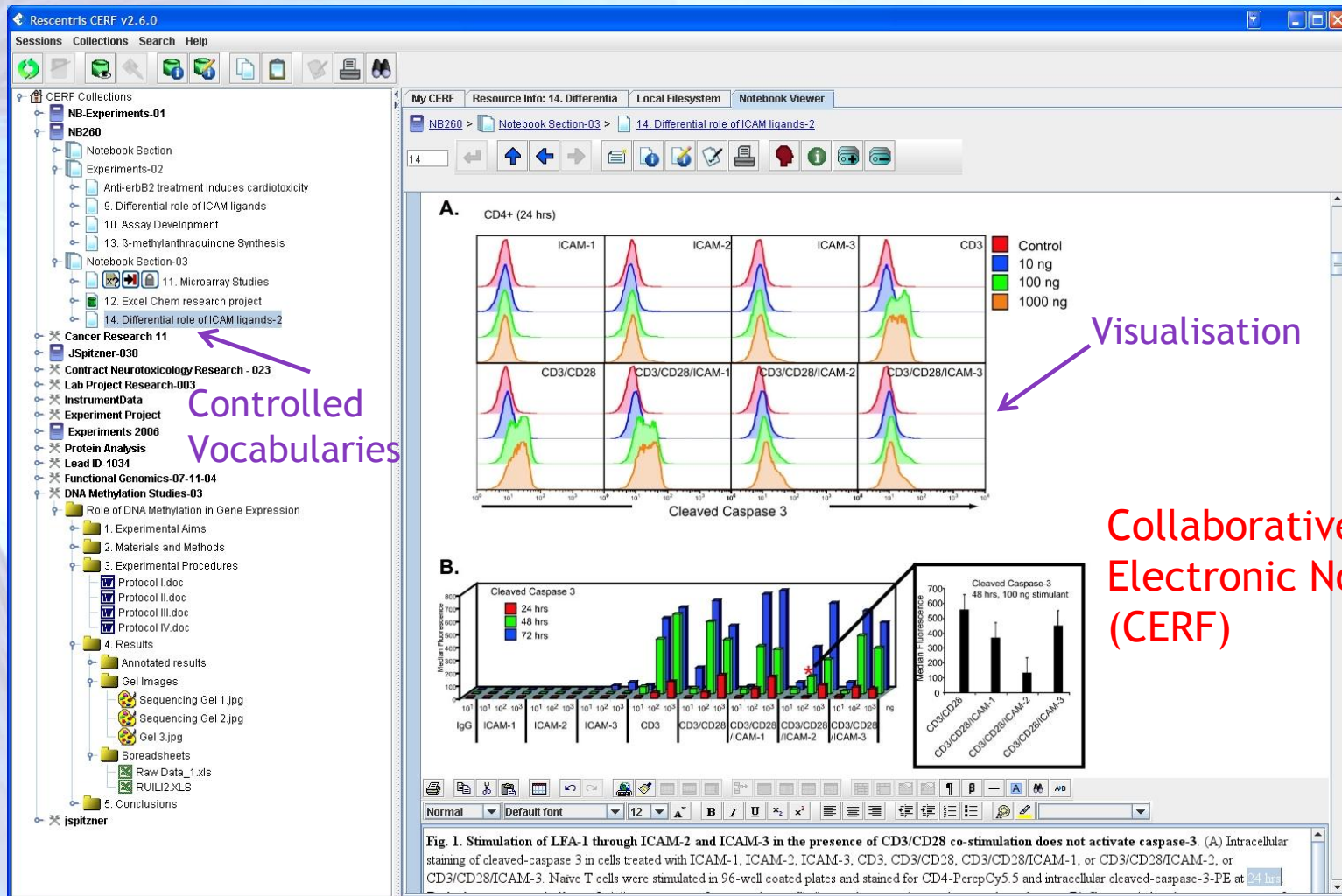
# Our Infrastructure Vision for ToxBank supporting all steps of Predictive Toxicology Research based on Alternative Testing methods

Users access compounds, biological materials, data and models for experimental planning and integrated analysis of experimental results



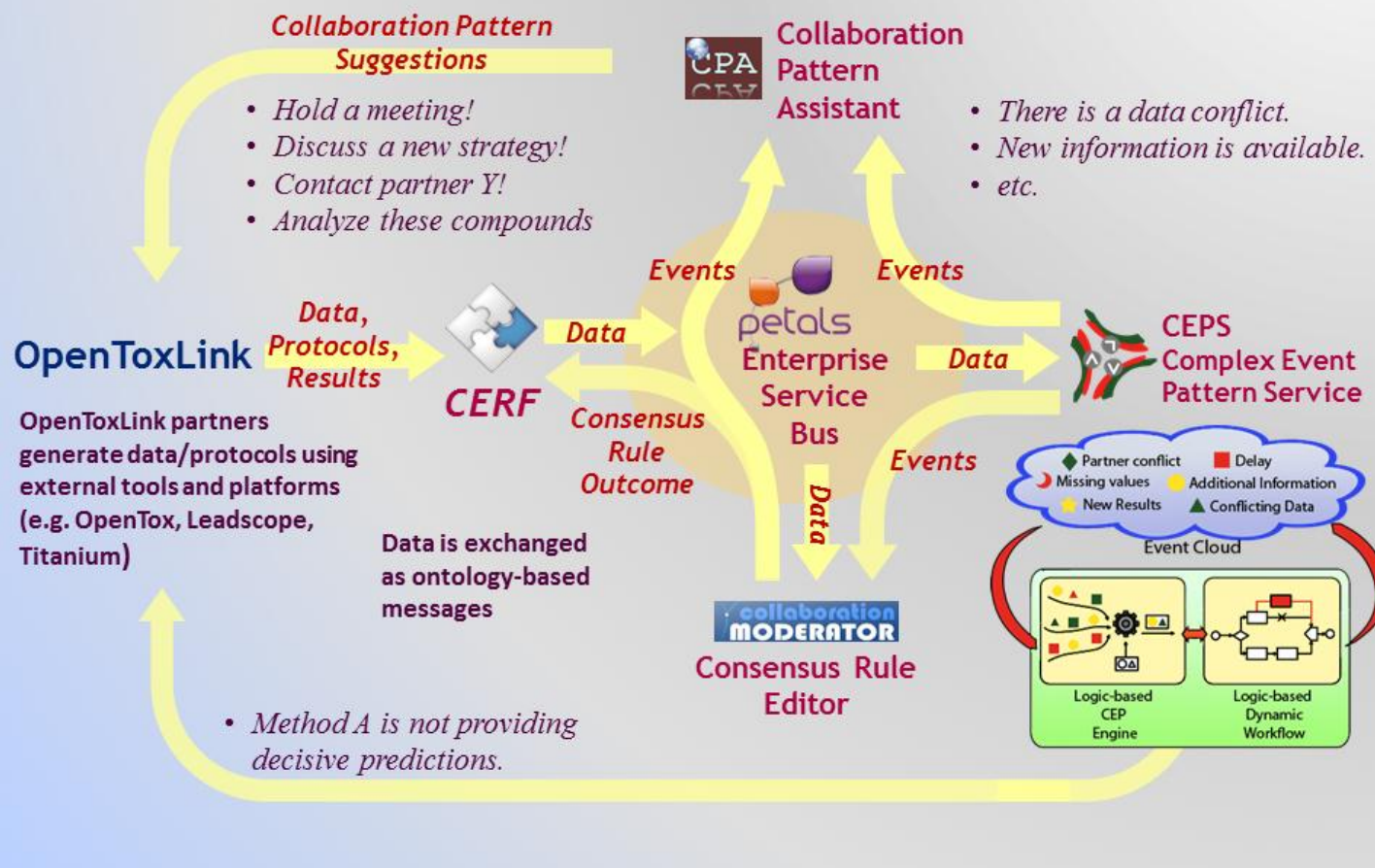


# Collaborative Research Framework Integration



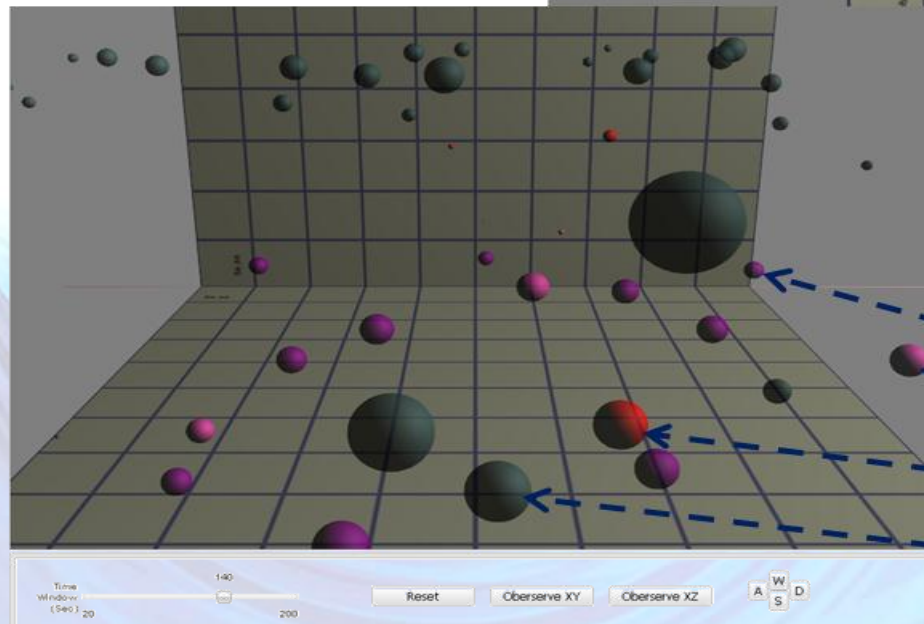
# Event Driven Collaboration Architecture

## OpenToxLink ICT Architecture



# Processing Complex Events Stream

Model 1	Model 2	Model 3		Assay 1	Assay 2	Assay 3	
1	0	1		-	-	-	



Showing	Symbol	PatternID	PatternName
<input checked="" type="checkbox"/>		carbon	Carbon
<input checked="" type="checkbox"/>		china	China

Symbol Pattern Name

- DOCK stopDOCK
- ADME stopADME
- TOX stopTOX



# Event Driven Weight of Evidence

CERF Client v4.0.0 - Logged in to Enterprise as jspitzner

Sessions Collections Bookmarks Search Tools Help

Project: Project-1001 Subject: Subject-1001 Compound Set: All Compound Sets Refresh Show Filters New Project New Subject New Compound Set New Compound Add Result

Results 1 to 100 of 197

Compound ID	Phone	VS	Dock	Dock 2	Binding Prediction Stoplight	QSAR ADME	QSPR ADME	ADME Prediction Stoplight	Binding + ADME Prediction Stoplight	Logic Based Tox	Limited Free Energy Tox	Toxicology Prediction Stoplight	Binding + ADME + Tox Prediction Stoplight	Saturation Binding Assay	Protein-DNA Binding Assay	Binding Assay Stoplight	In Vitro Toxicology Assay	In Vivo Toxicology Assay	Toxicology Assay Stoplight	Binding + Tox Assay Stoplight	Final Stoplight
UC0000353			0	0				0.0	-6.0999999												
UC0000862	1	1						-10.47	-10.8												
UC0000864	1	1						-10.2	-10.9												
UC0000884	1	1						-9.1400003	-10.6												
UC0000885	1	1						-9.1400003	-10.5												
UC0000886	1	1						-9.41	-10.6												
UC0000921	1	1						-10.91	-9.1000004												
UC0001349	1	1						-9.9799995	-11.2												
UC0001350	1	1						-9.96	-11.2												
UC0001500	1	1						-9.3299999	-9.3999996												
UC0001501	1	1						-9.5699997	-9.6000004												
UC0001623	1	1						-9.4899998	-9.1000004												
UC0001624	1	1						-9.4899998	-9.1000004												
UC0001699	1	1						-12.2	-10.9												
UC0001700	1	1						-9.9899998	-9.8000002												
UC0001702	1	1						-13.37	-9.6000004												
UC0001703	1	1						-10.61	-10.7												
UC0001743	1	1						-9.29	-9.1000004												
UC0001775	1	1						-9.7700005	-9.1000004												
UC0001875	1	1						-9.84	-9.2												
UC0001987	1	1						-9.7700005	-9.1999998												
UC0002838	1	1						-9.1999998	-9.8999996												
UC0002854	1	1						-10.09	-10.0												
UC0003266	1	1						-9.4799995	-9.8000002												
UC0003454	1	1						-9.1899996	-10.0												
UC0003835	1	1						-9.1000004	-9.8000002												
UC0003867	1	1						-10.25	-9.3999996												
UC0003923	1	1						-9.7200003	-9.8000002												
UC0003941	1	1						-10.52	-9.3000002												
UC0003973	1	1						-9.3100004	-9.1999998												

Previous Next Results per page: 100

Aggregate Resource

☒ Project ☐ Subject ☐ Compound Set ☐ Compound

Title: Project-1001

Status ?

Edit Status: Versionable

Owner: jspitzner

My Role: Notebook Creator

Closed: No

Checked Out: No

Visibility: Shared

Id: 26203 (Federation: 43214, Server: 801)

Metadata ?

Title: Project-1001

Submission/Modification

Resource Type: Drug Design Project

Creation Date: Oct 21, 2010 2:57:10 PM

Last Update: Oct 21, 2010 2:57:10 PM

Contributor: Jeff Spitzner

Relations and Annotations ?

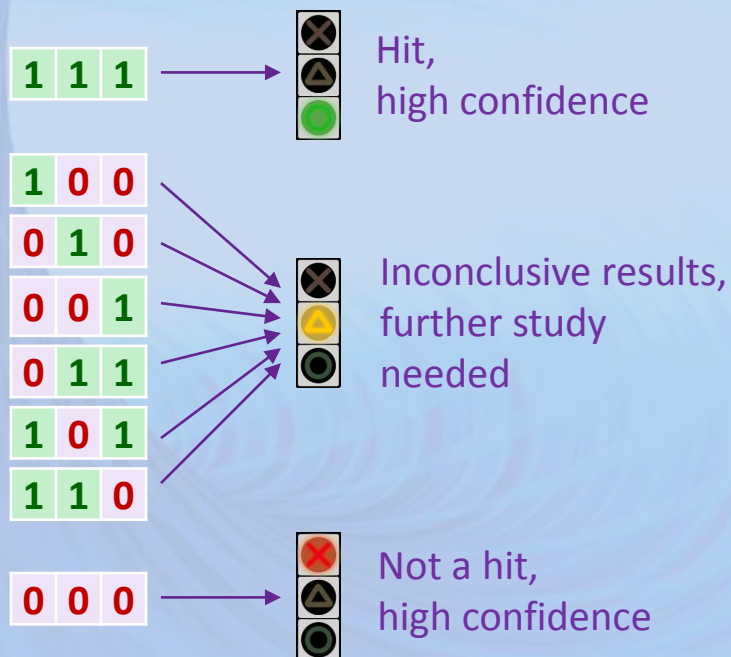
☐ Comment ☐ Tag

# Event Driven Weight of Evidence

collaboration  
**MODERATOR**

Consensus Rule  
Editor

## Recommendation Rules:



	Model 1	Model 2	Model 3		Assay 1	Assay 2	Assay 3	
	1	0	1		-	-	-	

**Synergy**

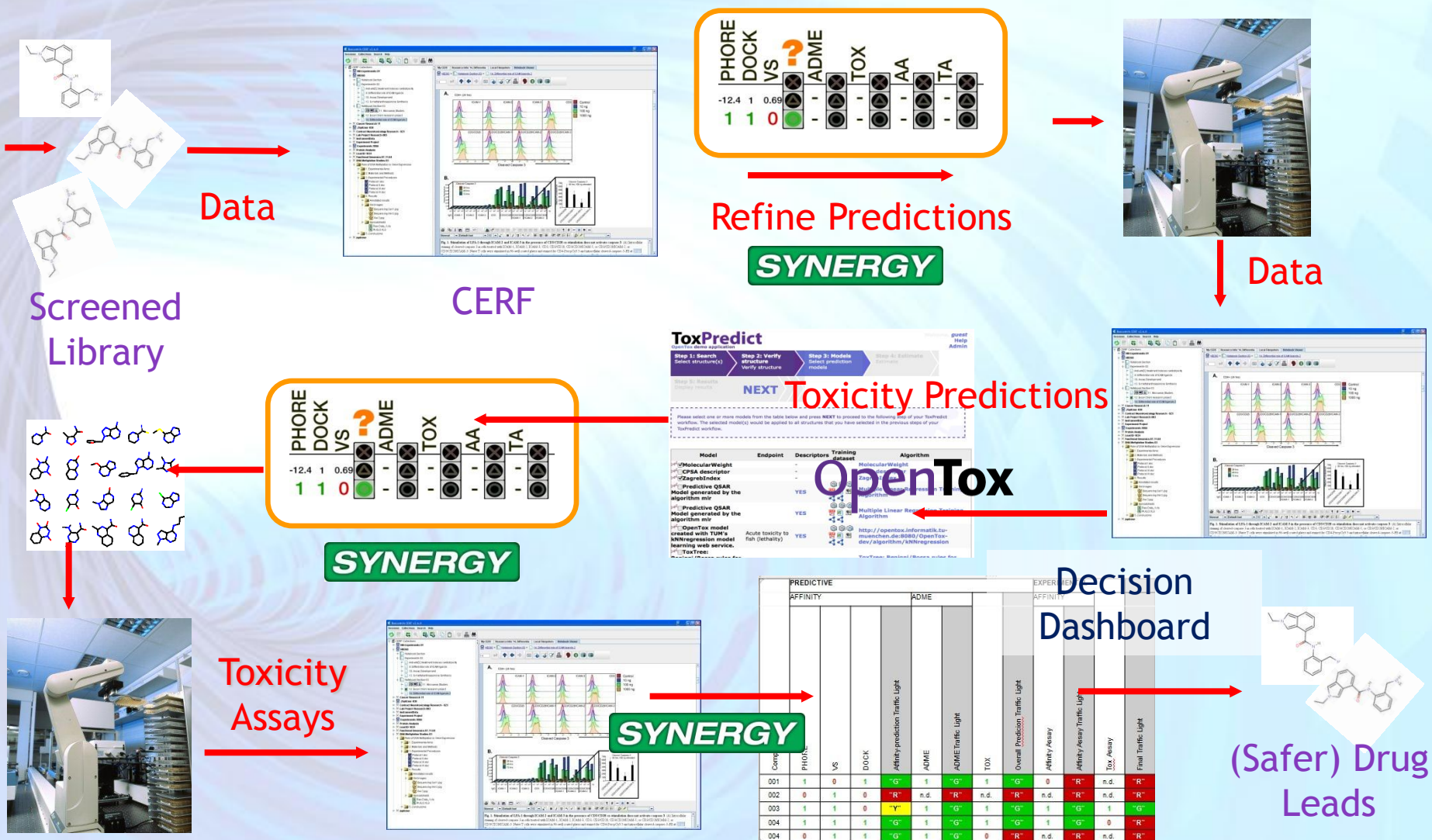
	Model 1	Model 2	Model 3		Assay 1	Assay 2	Assay 3	
	1	0	1		-	-	-	

**OpenTox**





# Synergy Drug Design Collaboration Pilot



# What are the benefits of OpenTox?

Through using components, open source and open, standard interfaces the OpenTox architecture enables:

- numerous applications in (Q)SAR, Read Across, reporting, metabolism prediction, etc. to be developed
- reliable integration of multiple applications
- easier retrieval and integration of data from multiple resources
- advancement of the *in silico* predictive toxicology field
- new improved algorithms and models to be built which can be used for REACH endpoints
- The integration of (Q)SAR into workflows eg. for Weight of Evidence



# OpenTox Workshop with 90 Participants in Rhodes

**Rhodos, Greece**

**Sept. 2010**

**OpenTox 3rd meeting**

**EuroQSAR 2010**

**Ca. 90 participants engaged in  
using a variety of OpenTox  
applications**



# OpenTox InterAction Meeting

## Innovation in Predictive Toxicology

Modeling, Applications, REACH, Risk Assessment

9-12 August, 2011

Technical University of Munich, Germany

---

**Registration:** Free but limited to 100 attendees

<https://www.surveymonkey.com/s/opentox2011>

**Abstracts by:** 30 April 2011, **Posters:** 30 June 2011

**Bursary Award:** three abstracts will be selected for travel & accommodation awards (by 30 April 2011)

**More Information at:**

[www.opentox.org/meet/opentox2011](http://www.opentox.org/meet/opentox2011)





# OpenTox and REACH Workshop

9 August 2011, Technical University Munich

## Practical Workshop on how OpenTox satisfies REACH requirements:

- 1 Supporting Information Gathering & Evaluation
- 2 Data Sourcing and Integration
- 3 (Q)SARS (including Applicability Domain, Validation, QMRF, QPRF reporting, Reliability, Confidence)
- 4 Categories and Read Across
- 5 Weight of Evidence
- 6 Predicting Metabolites
- 7 Achieving Harmonisation through Ontologies

An extensive manuscript will be prepared and submitted for publication.



# Collaborating Partners

In Silico Toxicology,  
Switzerland

Douglas Connect,  
Switzerland  
(Coordinator)

Albert Ludwigs University  
Freiburg, Germany

Ideaconsult,  
Bulgaria

Istituto Superiore  
di Sanità, Italy

Technical University  
of Munich, Germany



National Technical  
University of Athens,  
Greece

Fraunhofer Institute  
for Toxicology &  
Experimental Medicine,  
Germany

David Gallagher, UK

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

Seascope Learning &  
JNU, India

# Our Funding Support...

OpenTox - An Open Source Predictive Toxicology Framework, [www.opentox.org](http://www.opentox.org), is funded under the EU Seventh Framework Program: HEALTH-2007-1.3-3 Promotion, development, validation, acceptance and implementation of QSARs (Quantitative Structure-Activity Relationships) for toxicology  
Project Reference Number Health-F5-2008-200787 (2008-2011).

For more information, visit

[www.opentox.org](http://www.opentox.org)

Contact me:

[barry.hardy -\(at\)- douglasconnect.com](mailto:barry.hardy-(at)-douglasconnect.com)

[twitter.com/barryhardy](https://twitter.com/barryhardy)

[barryhardy.blogs.com](http://barryhardy.blogs.com)