OpenTox Services and Applications

Barry Hardy (Douglas Connect)

AXLR8 Meeting
Berlin, Germany
23 May 2011

OpenTox

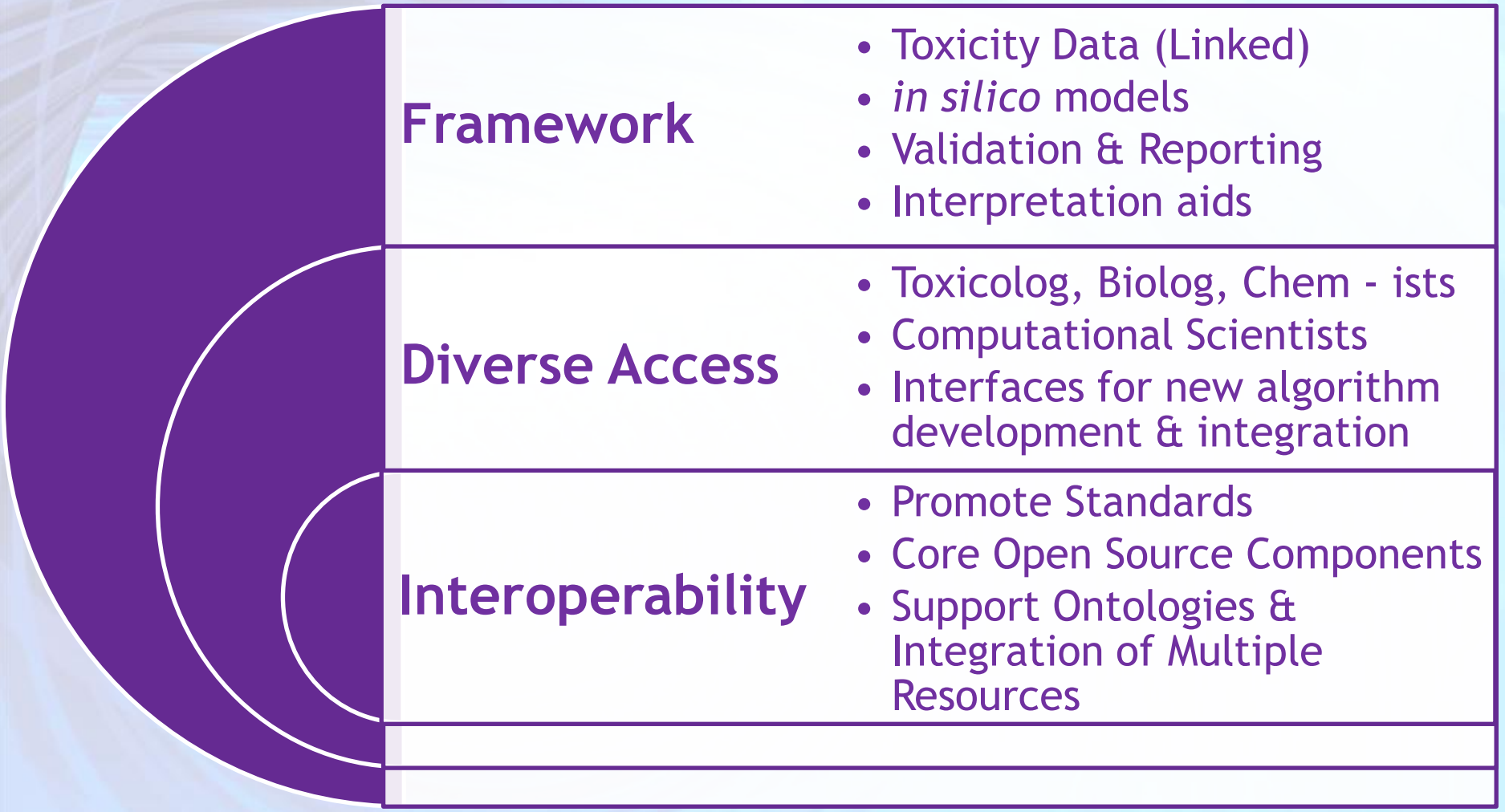
SEVENTH FRAMEWORK PROGRAMME

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.

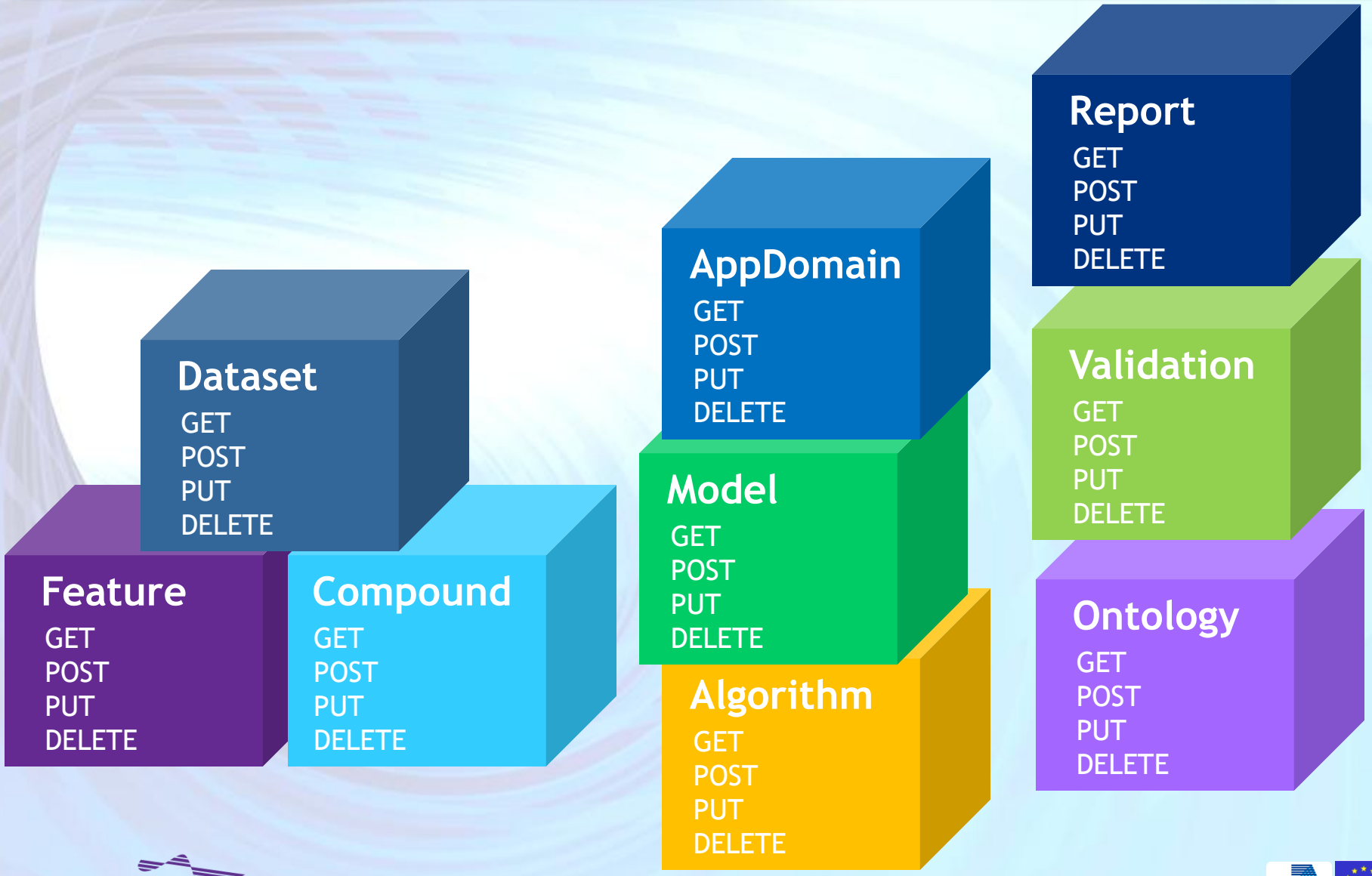
OpenTox is an Integrating Framework



OpenTox

SEVENTH FRAMEWORK PROGRAMME

Overview of OpenTox Application Programming Interfaces



OpenTox

The OpenTox Framework (reported last year)

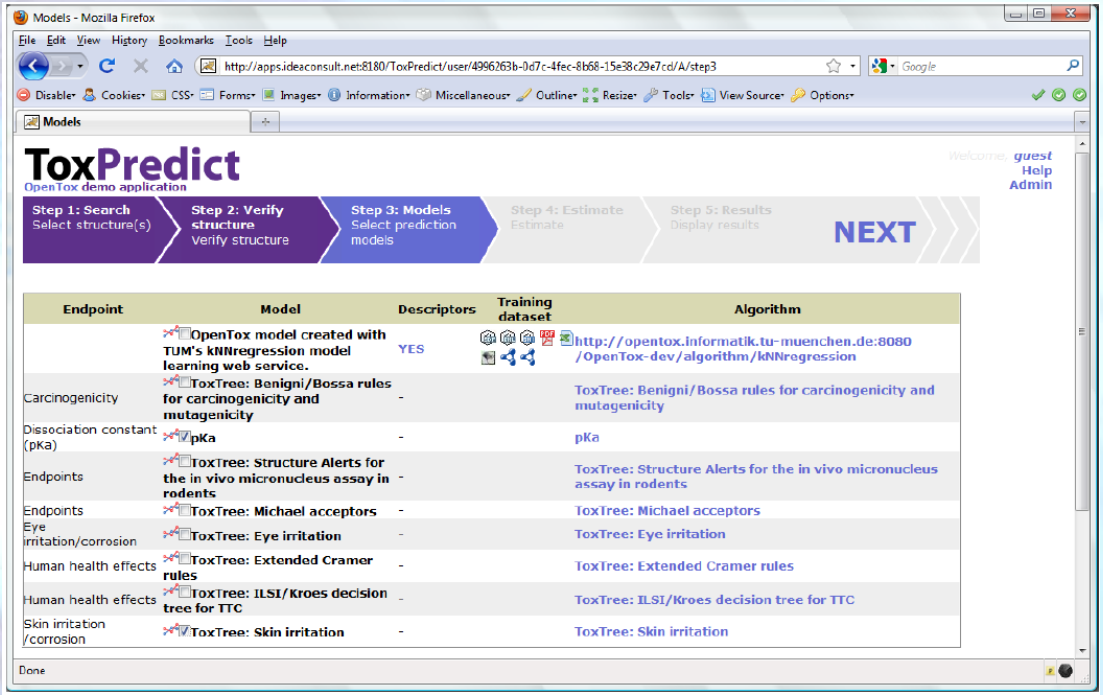
Collaborative development of predictive toxicology applications
Journal of Cheminformatics 2010, 2:7 doi:10.1186/1758-2946-2-7

Barry Hardy, Nicki Douglas, Christoph Helma, Micha Rautenberg, Nina Jeliazkova, Vedrin Jeliazkov, 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 Gloriozova, 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

OpenTox

What you can do with it …



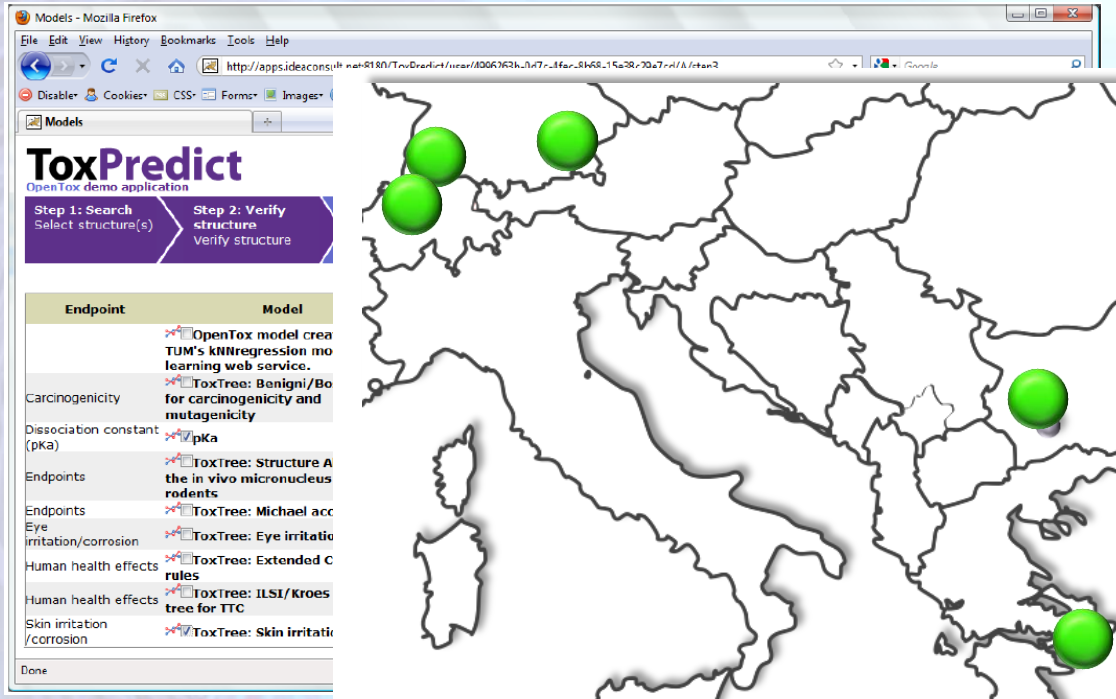
Simple building of predictive toxicology applications based on well-established methods and databases

OpenTox

Developed by Ideaconsult

SEVENTH FRAMEWORK PROGRAMME

What you can do with it ...



Simple bui

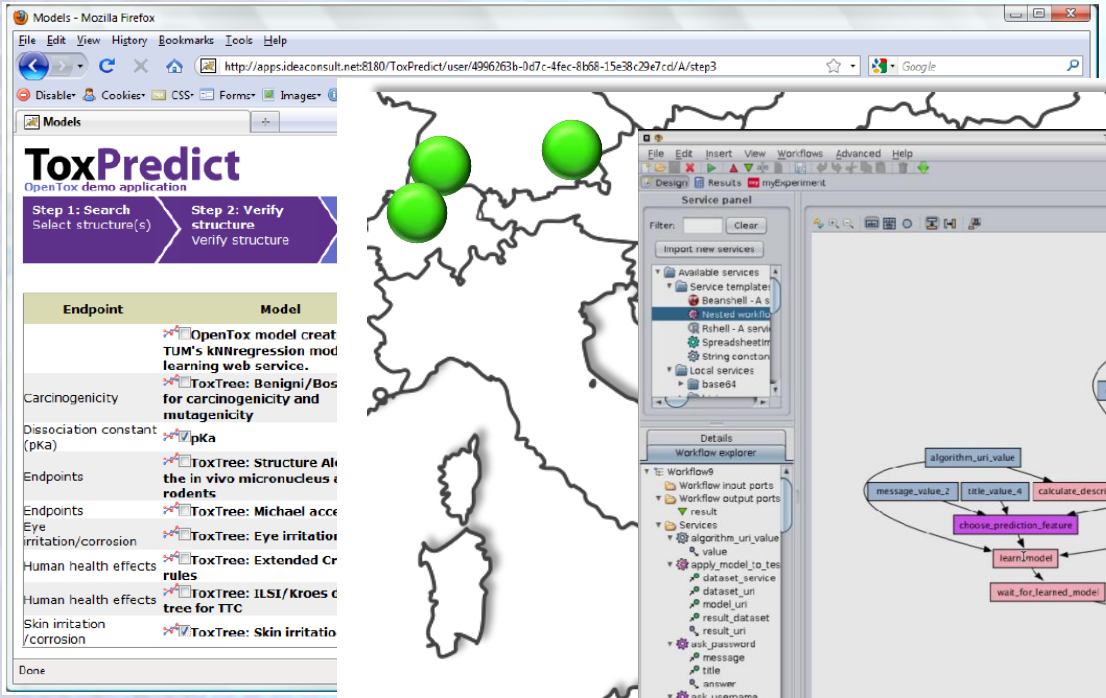
applicatio

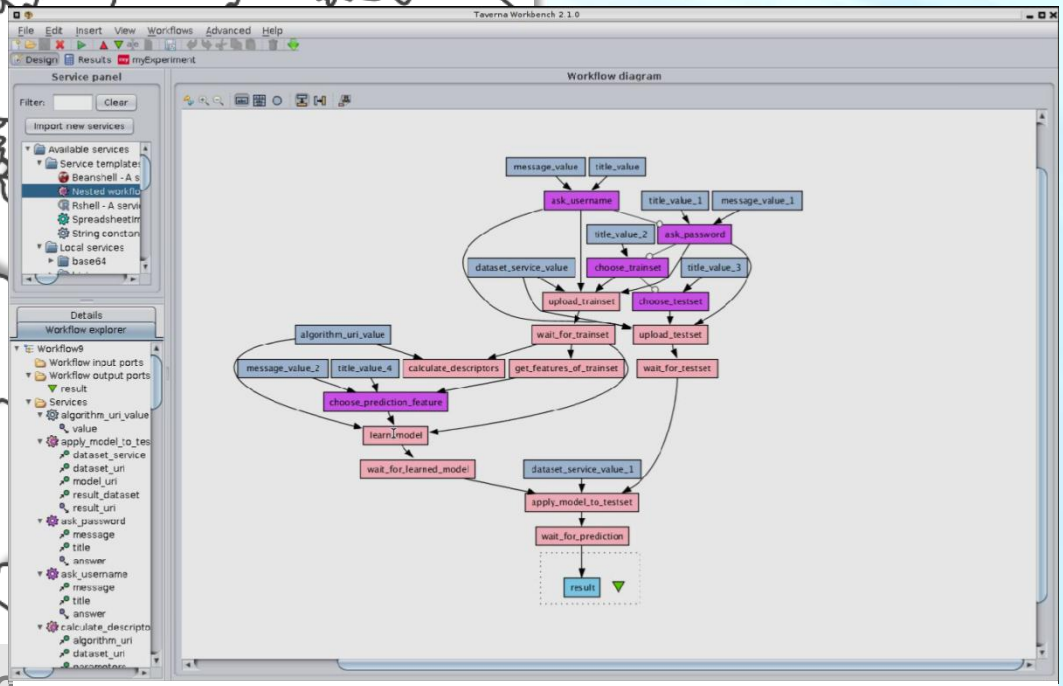
methods a

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

OpenTox

What you can do with it …





Simple buildi
applications
methods and

Distributed a
wide range o
methods

Integration into workflow systems
for computational biology

OpenTox

Need for Applications for REACH

OpenTox

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

OpenTox

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)

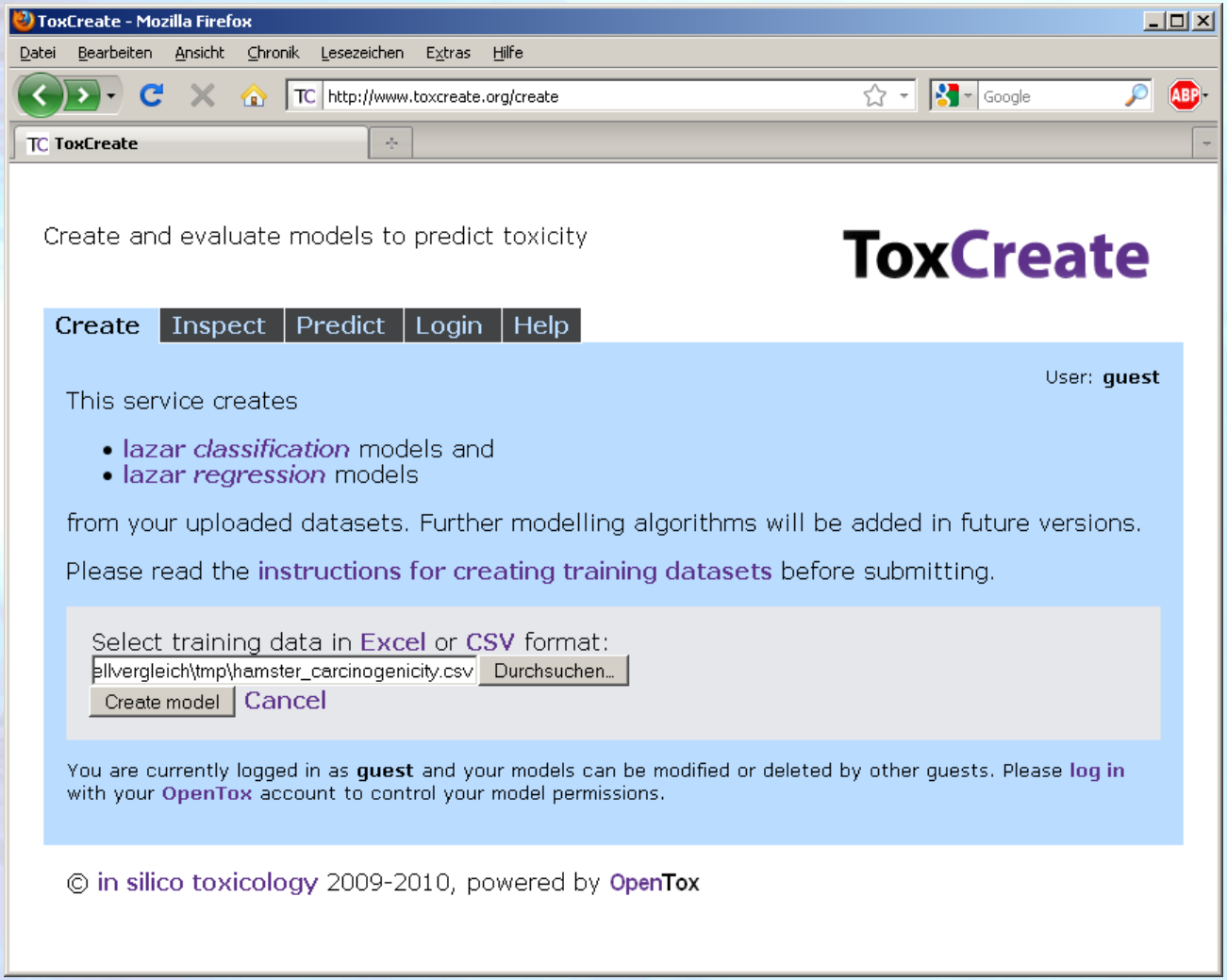
(Q)SARs & REACH requirements

(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.

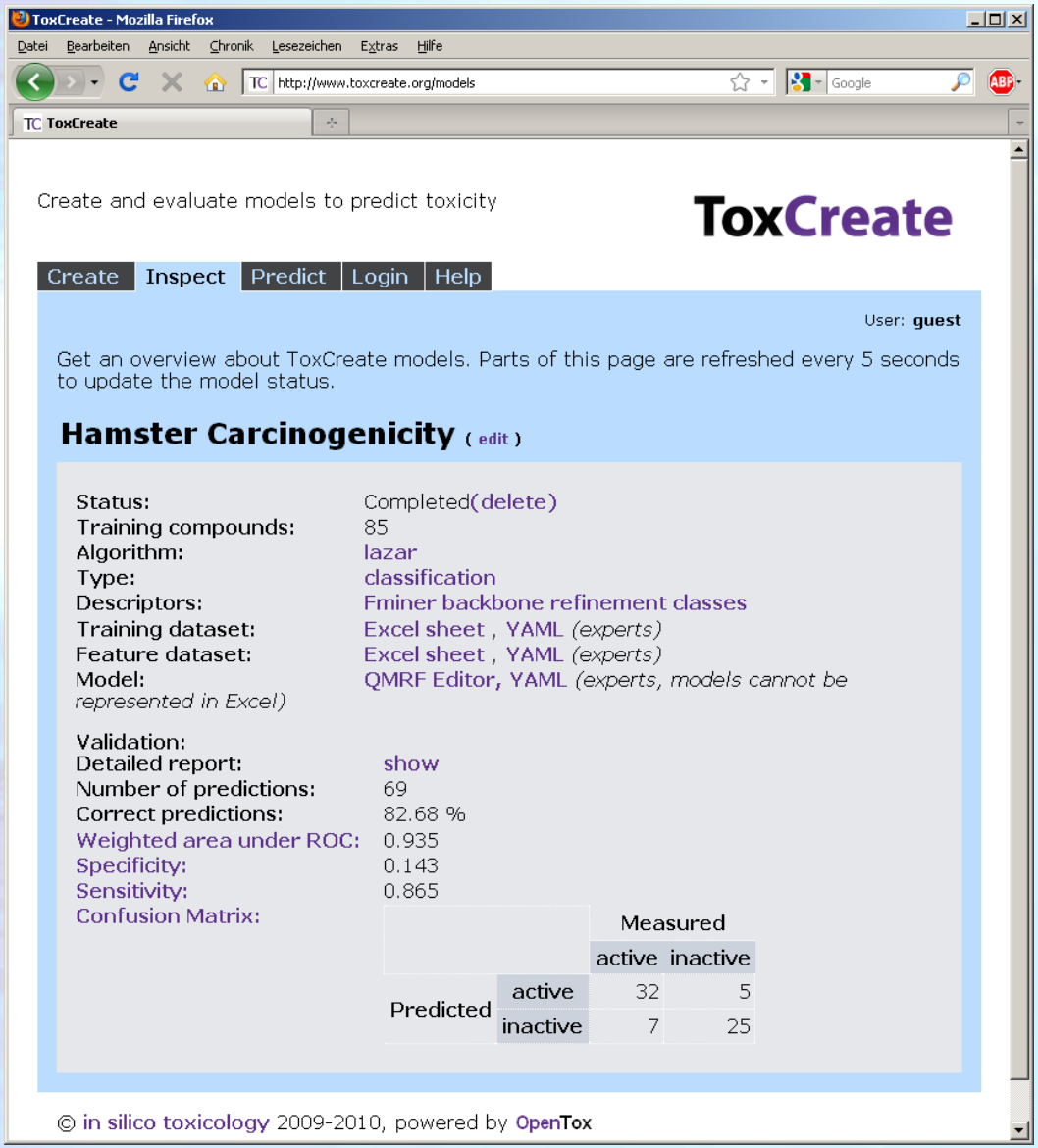
ToxCreate - (Q)SAR Model Building application



Developed by In Silico Toxicology

OpenTox

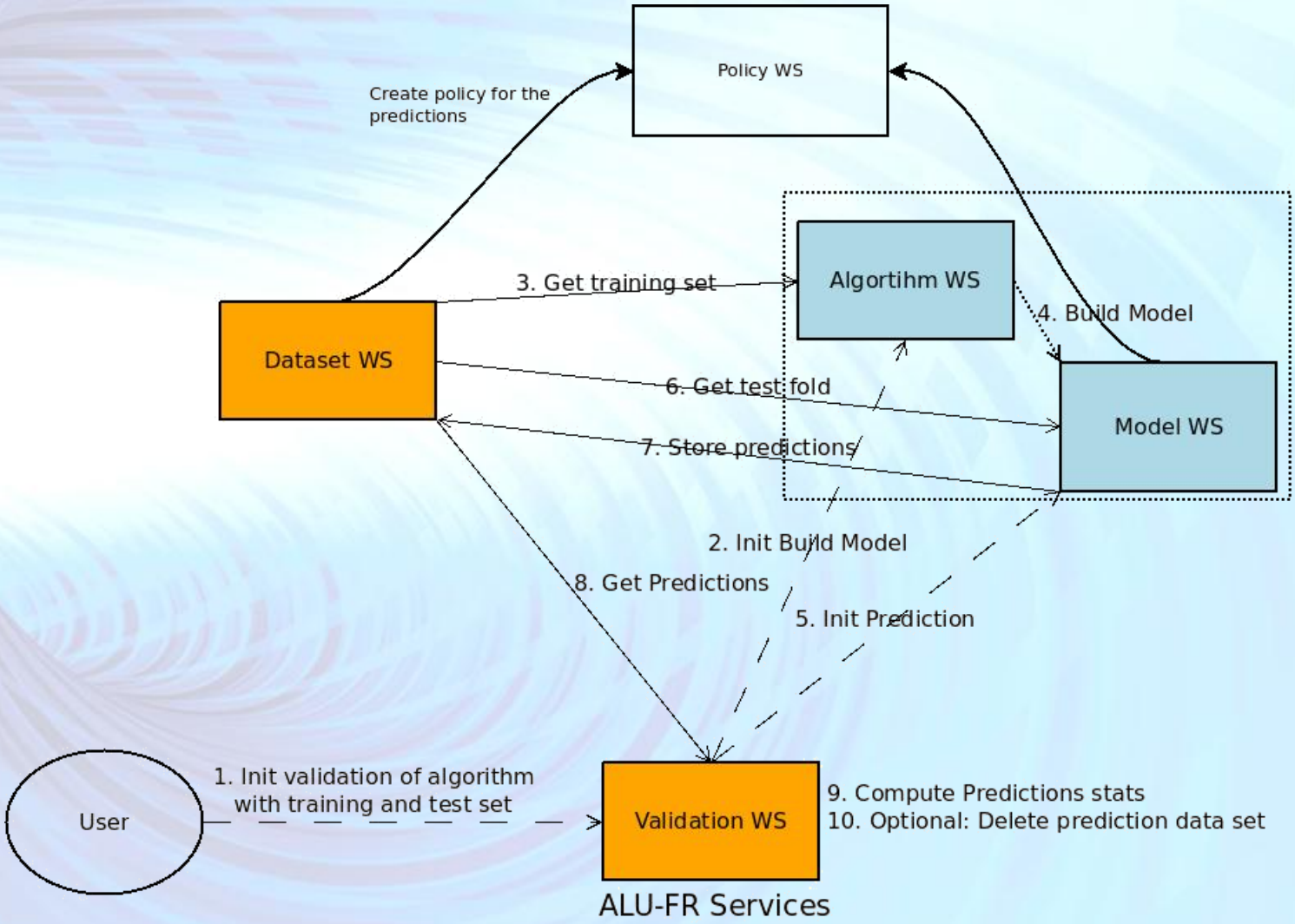
ToxCreate – (Q)SAR Model Results



OpenTox

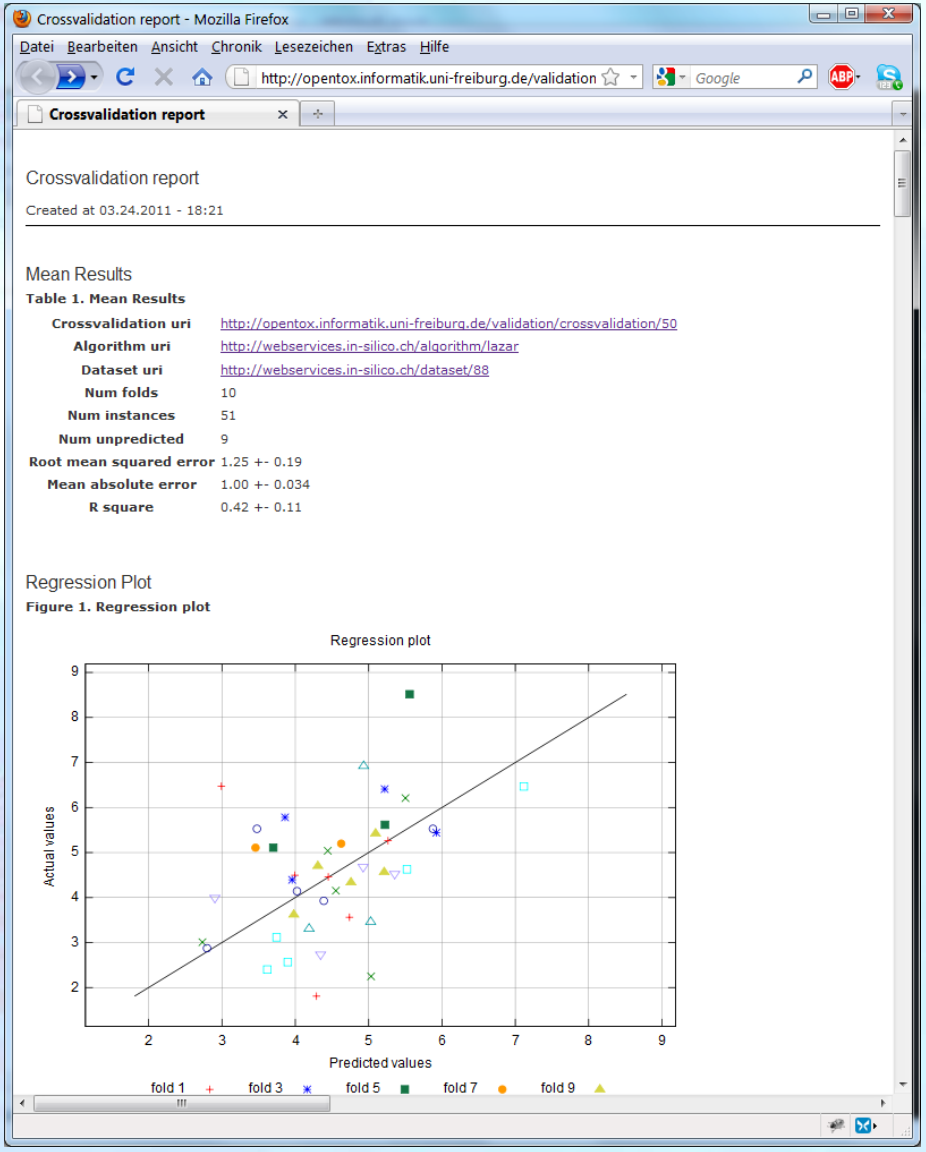
	OECD Principle	OpenTox addresses Validation Principles by...
1	Defined Endpoint	providing a unified source of well defined and documented toxicity data with a common vocabulary
2	Unambiguous Algorithm	providing transparent access to well documented models and algorithms as well as to the source code
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



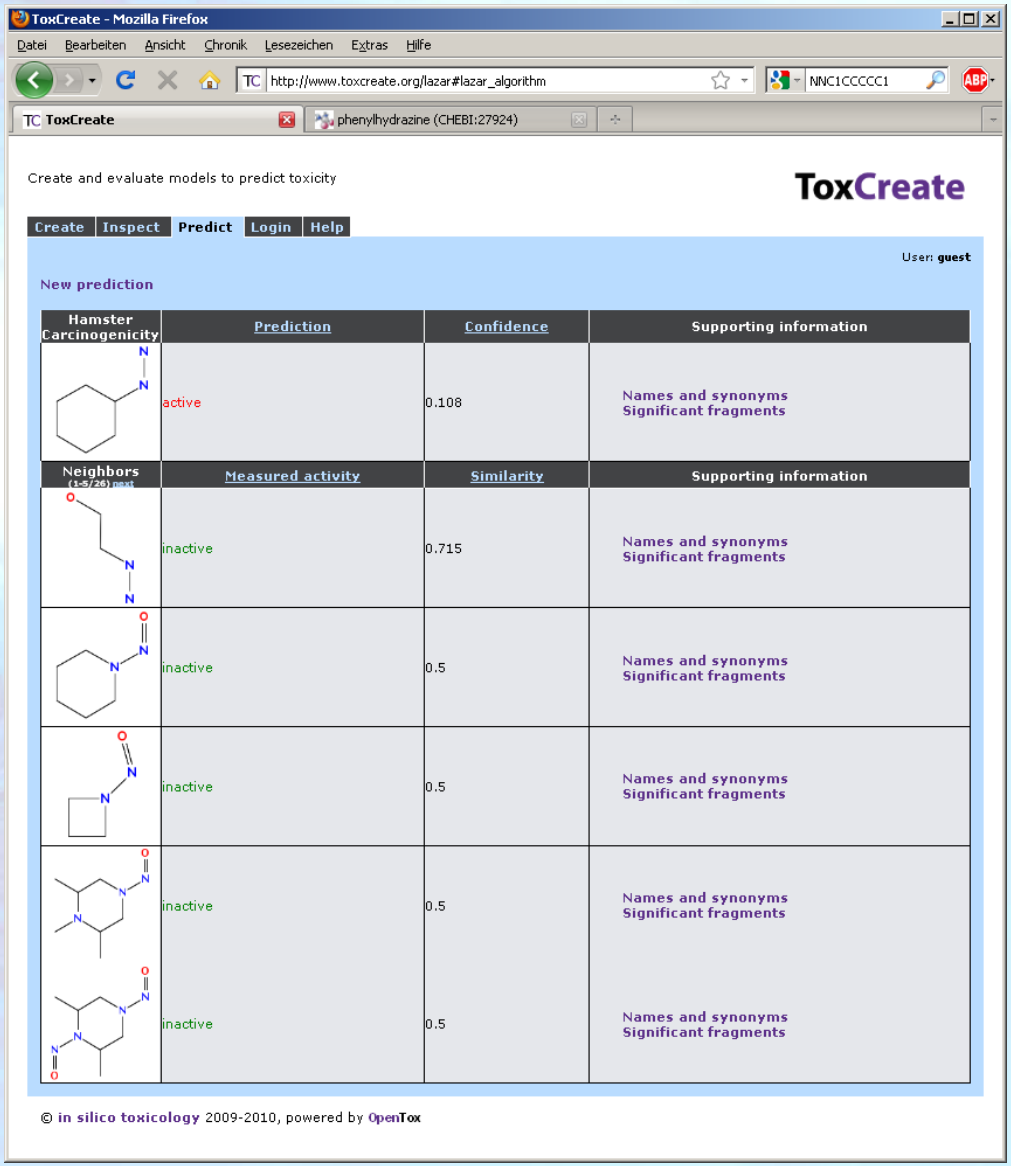
OpenTox

ToxCreate - linked to Validation Service



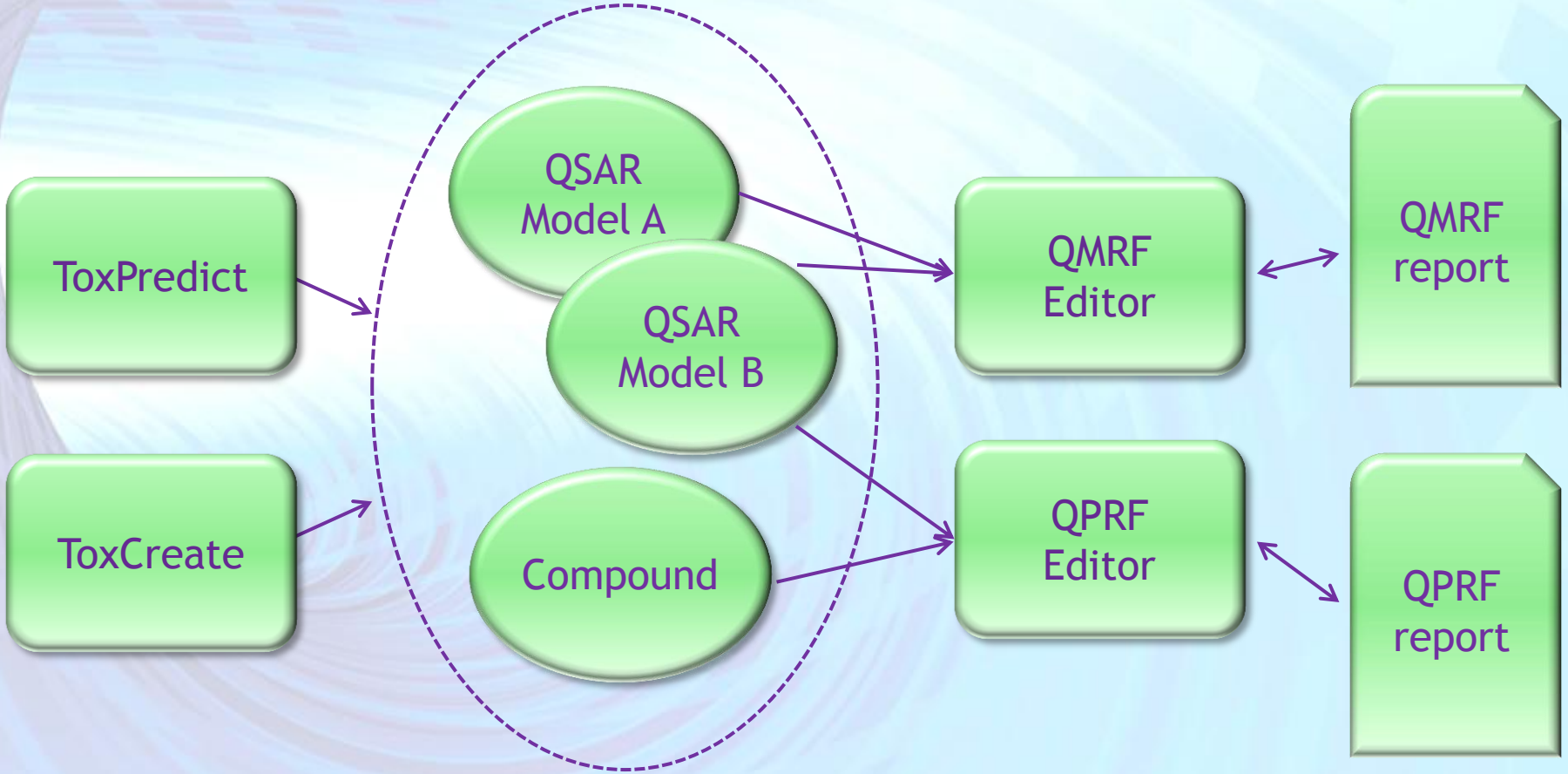
OpenTox

ToxCreate – Confidence, Supporting Information

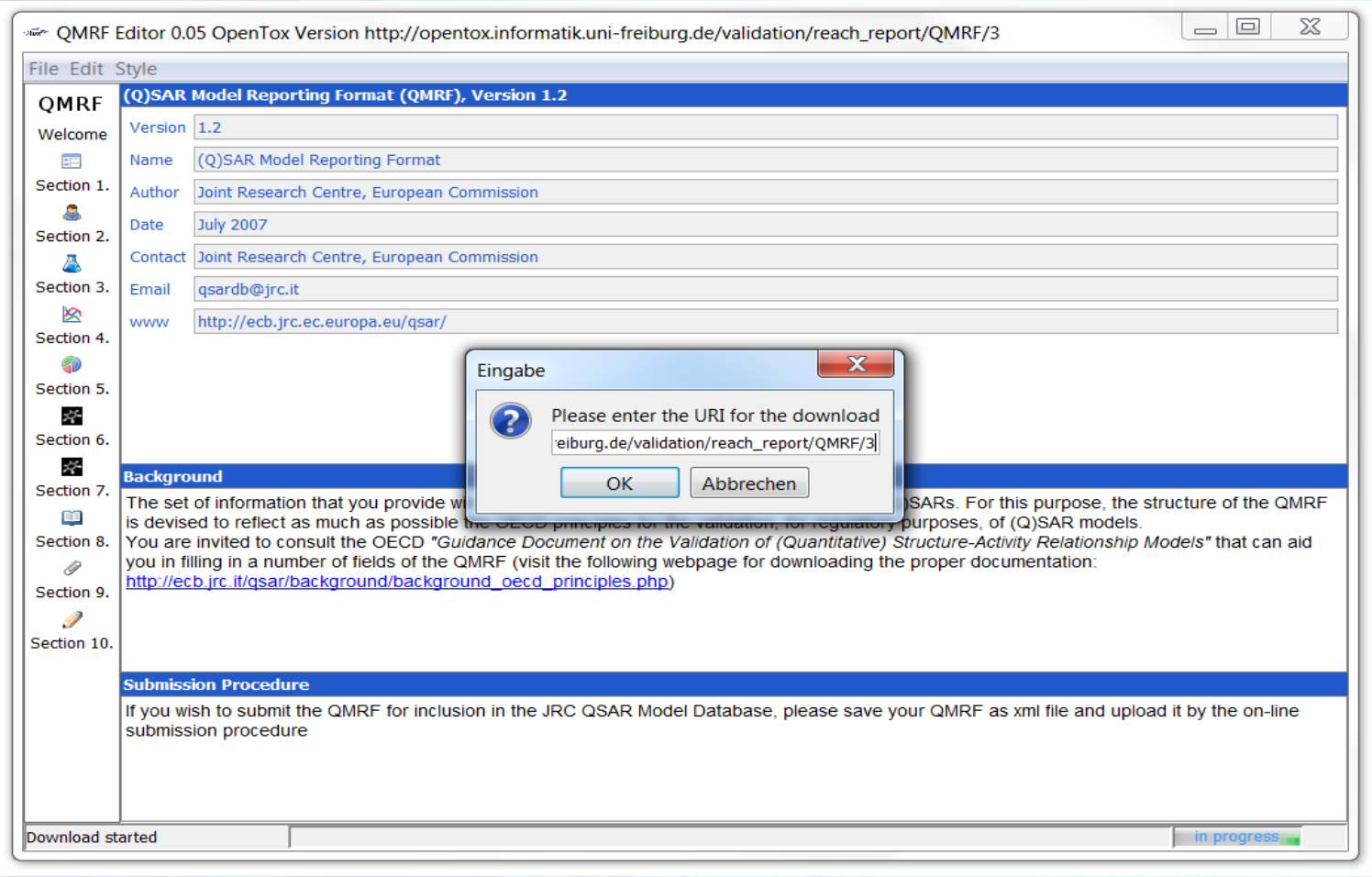


OpenTox

(Q)SARs - reporting in OpenTox

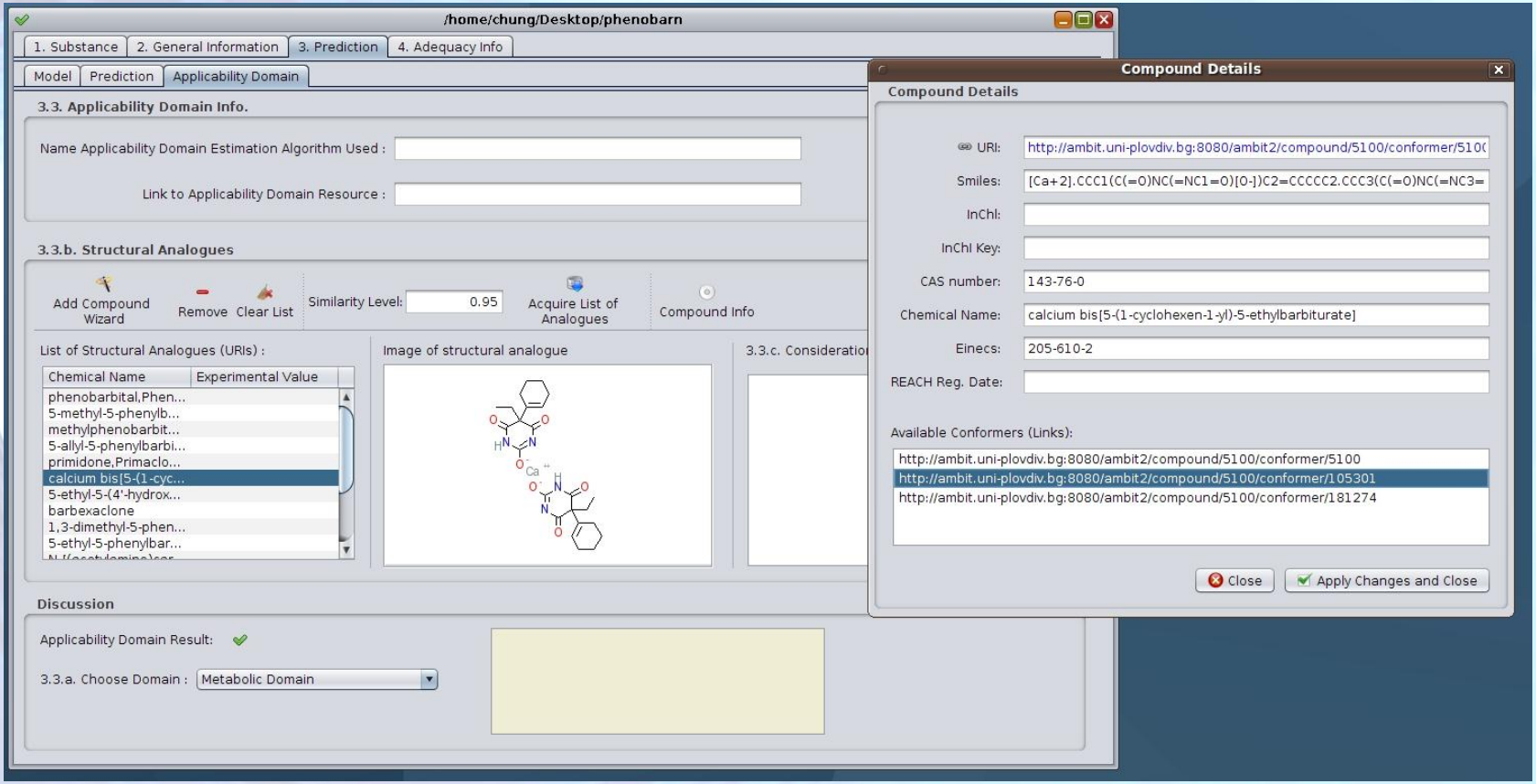


(Q)SARs - QMRF reporting in OpenTox



OpenTox

QPRF Reporting (Qedit)



Application by Pantelis Sopasakis (NTUA)

OpenTox

Metabolites

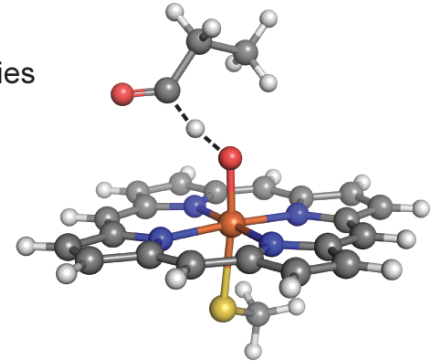
According to ECHA Guidance B, further investigation may be required for degradation products and metabolites if considered relevant for the chemical safety assessment, PBT assessment or classification and labeling.

SMARTCyp Service for Predicting Metabolites

Atom Reactivity Library

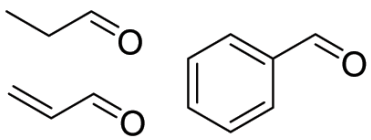
A. Calculate Quantum Chemical Reference Energies

Calculate transition state energies using density functional theory



B. Define SMARTS Rules

Group calculations by fragments and calculate average energies



[CX3H1](=O)[#6]

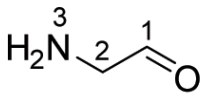
40.2 kJ/mol

SMARTCyp - developed by Patrik Rydberg, University of Copenhagen

www.farma.ku.dk/index.php/SMARTCyp/7990/0/

SMARTCyp

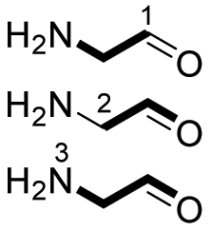
1. Assign Energies By SMARTS matching



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

2. Compute Accessibility Descriptor

$A_i = Maxbonds_i / Maxbonds_{all}$



$A_1 = 2 / 3 = 0.67$

$A_2 = 2 / 3 = 0.67$

$A_3 = 3 / 3 = 1.00$

3. Compute Score and Rank Atoms

Score, S = E - 8A
Lowest score gets rank 1

$S_1 = 40.2 - 8*0.67 = 34.84$

$S_2 = 39.8 - 8*0.67 = 34.44$

$S_3 = 54.1 - 8*1.00 = 46.10$

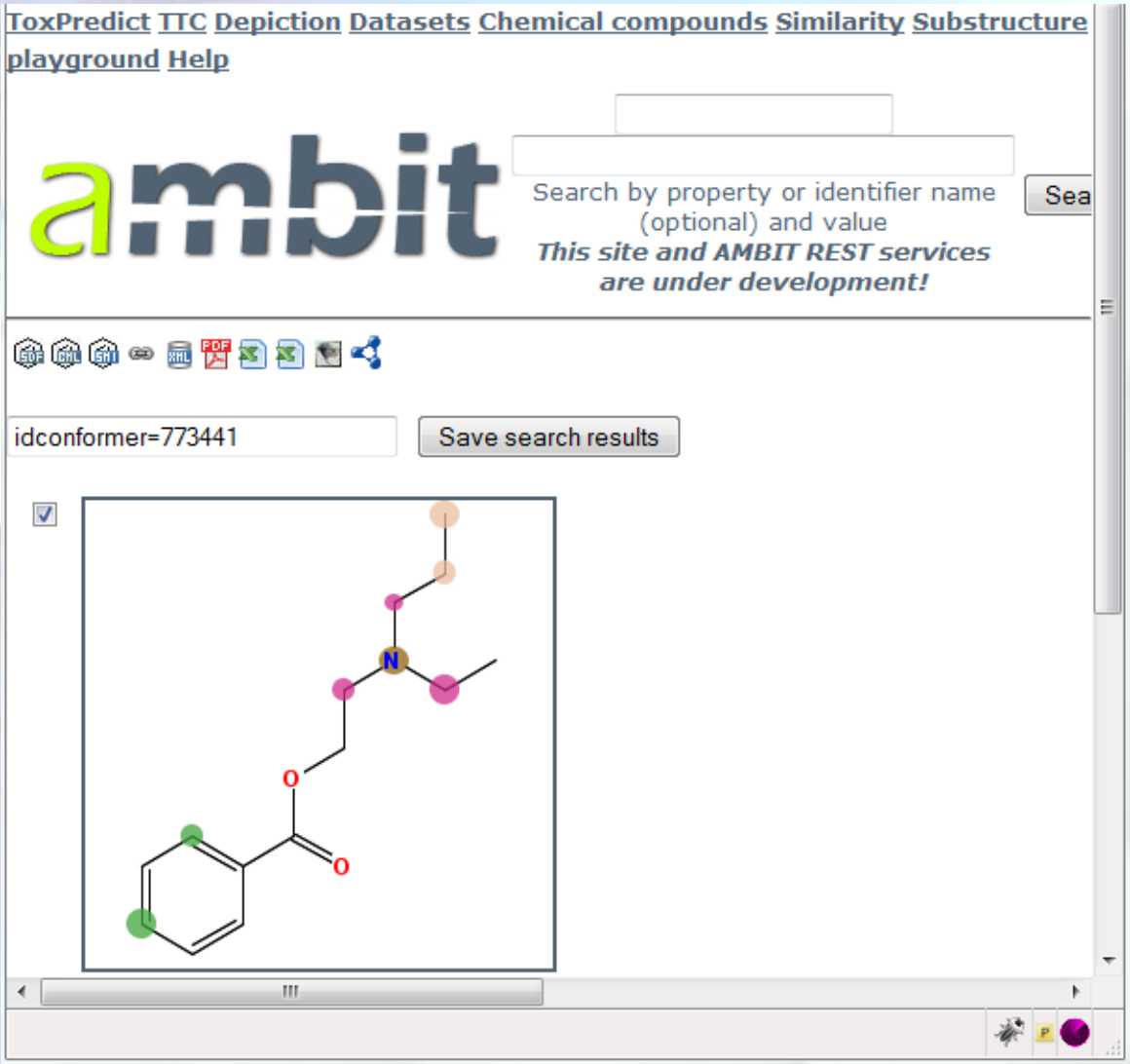
Atom 1 - Rank 2

Atom 2 - Rank 1

Atom 3 - Rank 3

OpenTox

Metabolites



Aliphatic Hydroxylation

Aromatic Hydroxylation

N-Dealkylation

N-Oxidation

O-Dealkylation

S-Oxidation

Developed by Ideaconsult

OpenTox

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.

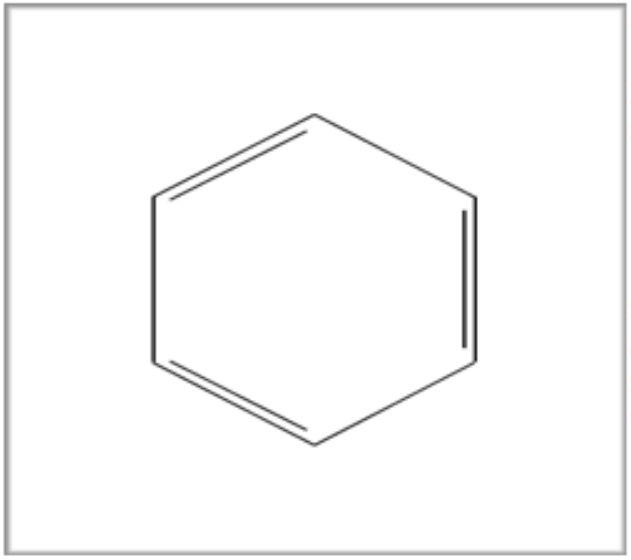
The Part B guidance states that this information may be obtained from a variety of sources including “databases or other sources in the open literature or accessible on the internet”.

ToxPredict

OpenTox demo application

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

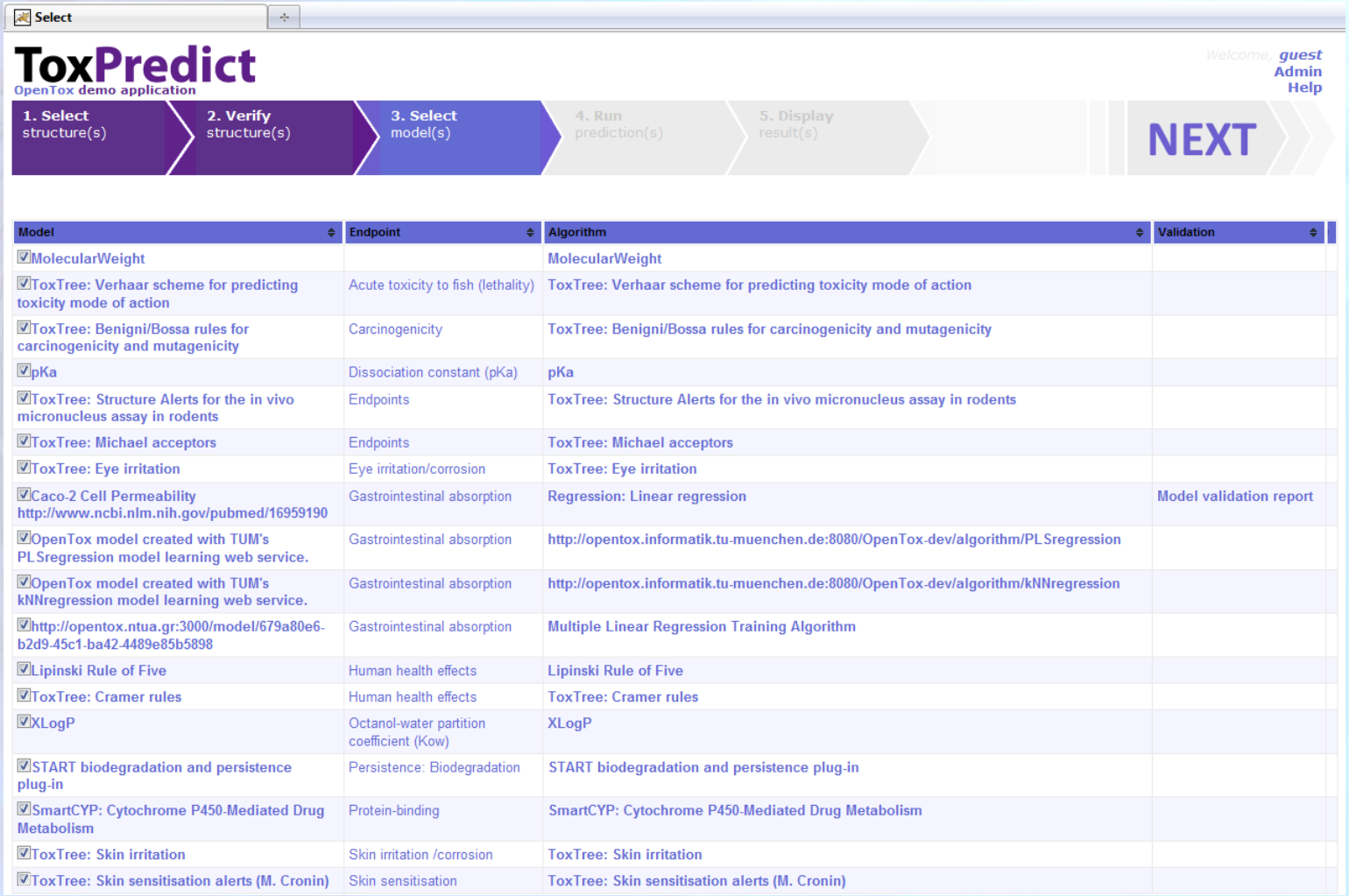


CAS RN	71-43-2
EINECS	200-753-7
IUPAC name	benzene
Synonym	(6)annulene; benzine; Benzol; Benzolene; bicarburet of hydrogen; carbon oil; Coal naphtha; cyclohexatriene; mineral naphtha; motor benzol; nitration benzene; Phene; Phenyl hydride; pyrobenzol.
Synonym	21742.0
Synonym	Benzene
Synonym	benzene
Quality label	OK

MolecularWeight MolecularWeight

MW	78.1112

Information Gathering & Evaluation Process

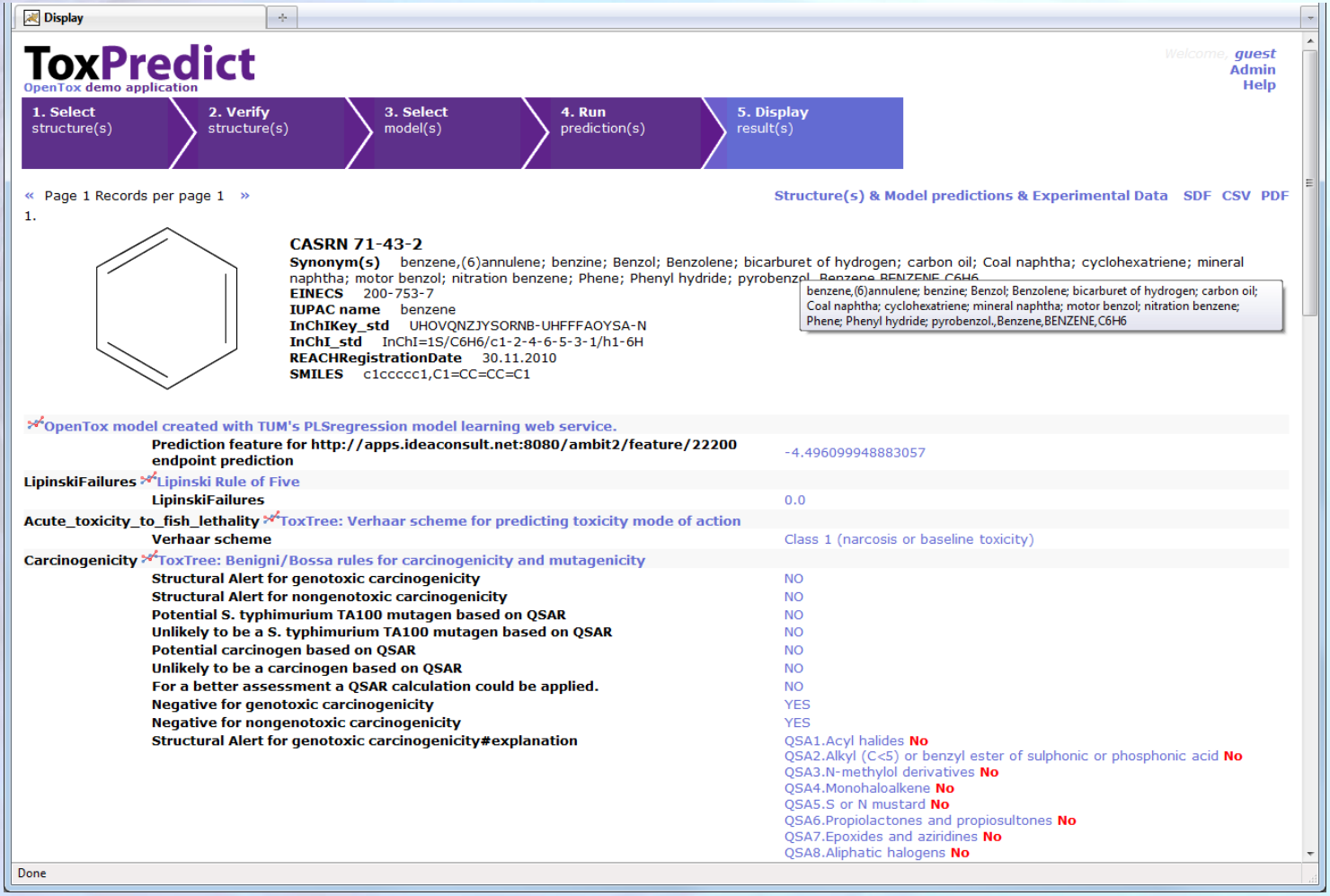


Model	Endpoint	Algorithm	Validation
MolecularWeight		MolecularWeight	
ToxTree: Verhaar scheme for predicting toxicity mode of action	Acute toxicity to fish (lethality)	ToxTree: Verhaar scheme for predicting toxicity mode of action	
ToxTree: Benigni/Bossa rules for carcinogenicity and mutagenicity	Carcinogenicity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

OpenTox

SEVENTH FRAMEWORK PROGRAMME

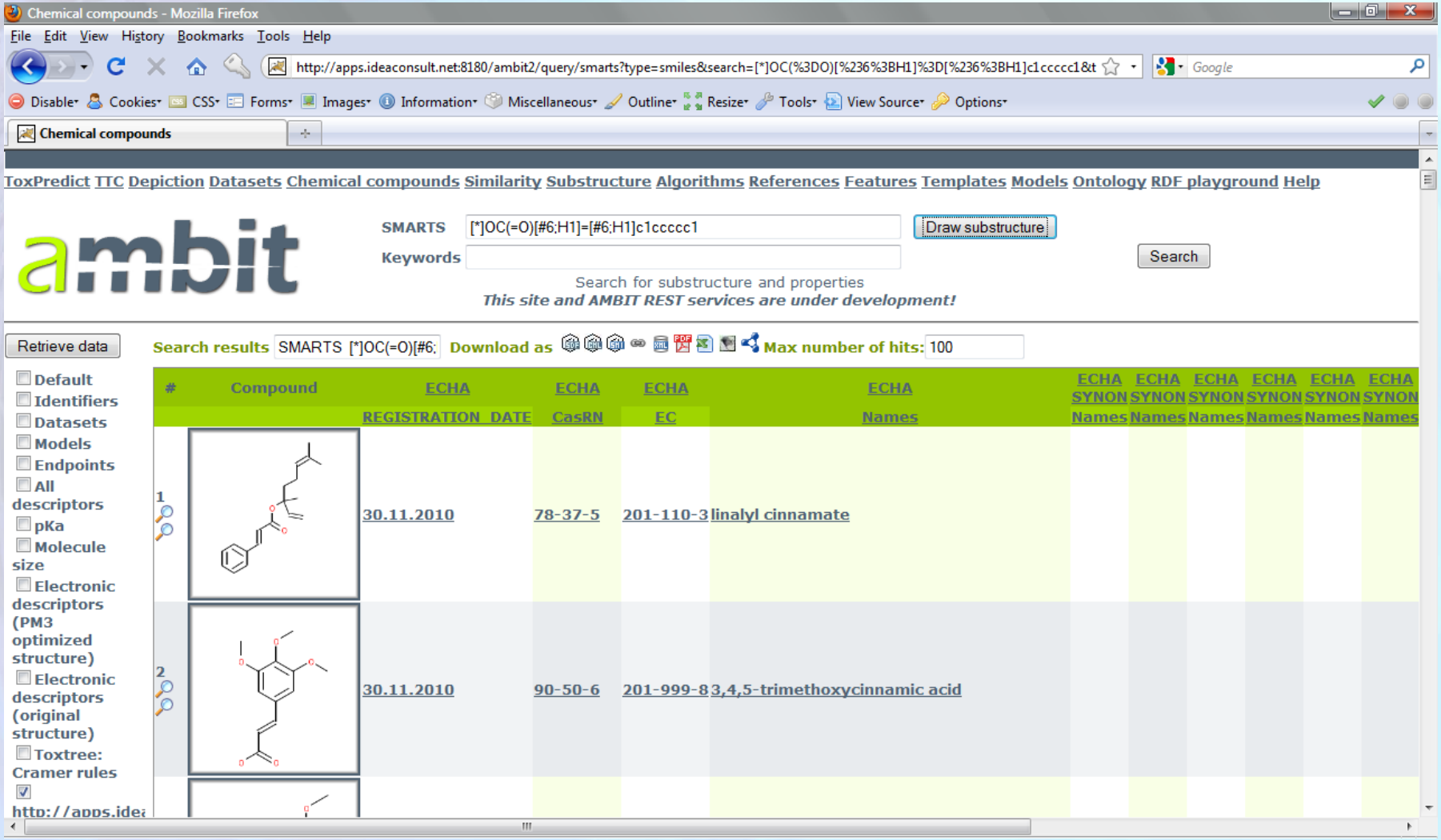
Information Gathering & Evaluation Process



OpenTox

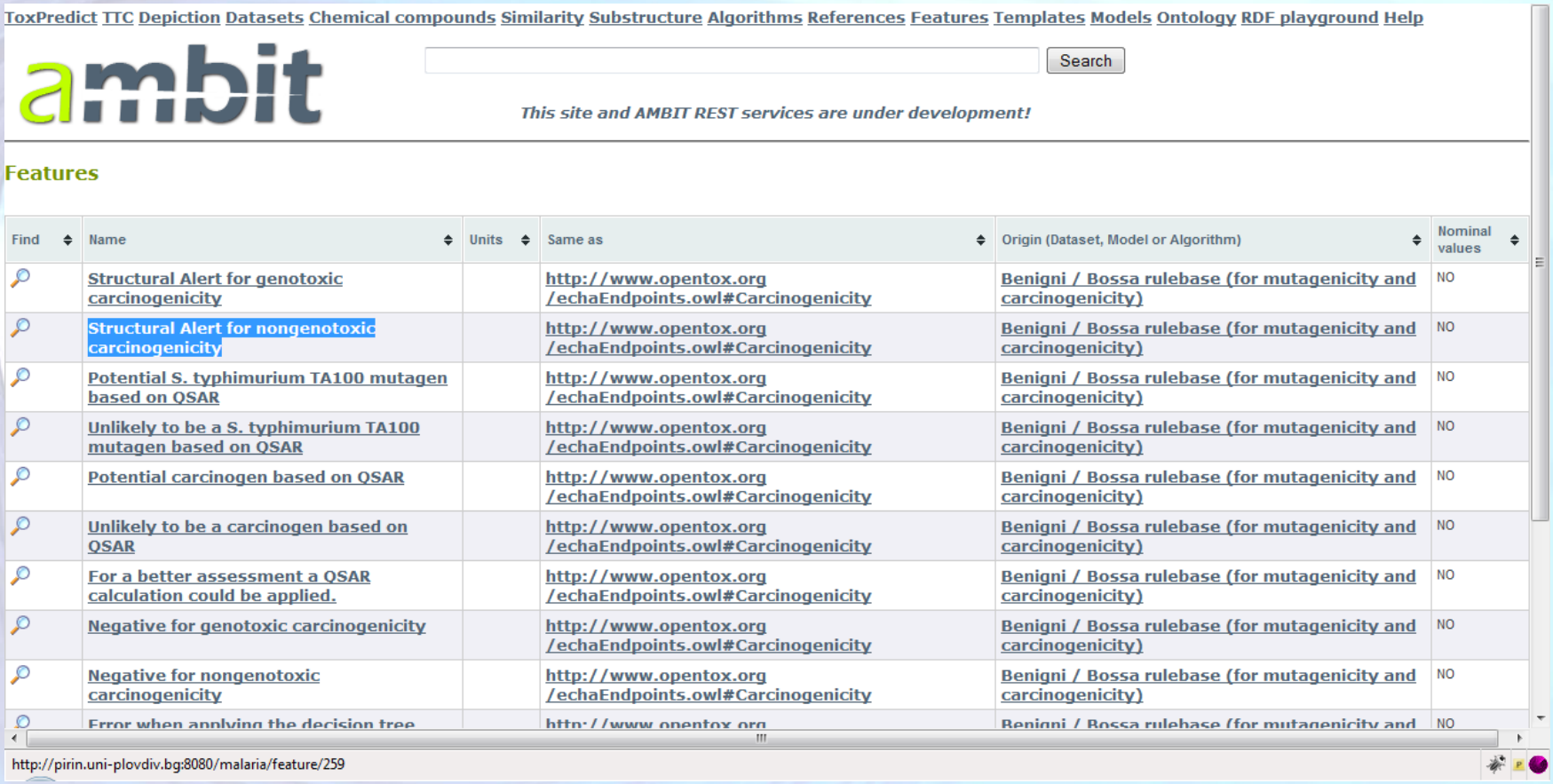
SEVENTH FRAMEWORK PROGRAMME

Information Gathering & Evaluation Process



OpenTox

Information Gathering & Evaluation Process



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

ambit

Search

This site and AMBIT REST services are under development!

Features

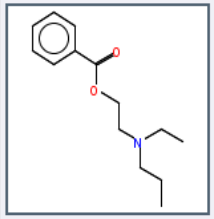
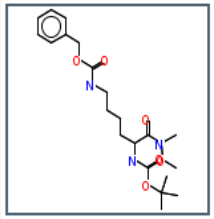
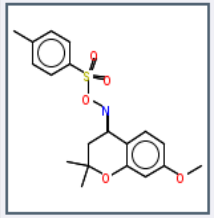
Find	Name	Units	Same as	Origin (Dataset, Model or Algorithm)	Nominal values
🔎	Structural Alert for genotoxic carcinogenicity		http://www.opentox.org/echaEndpoints.owl#Carcinogenicity	Benigni / Bossa rulebase (for mutagenicity and carcinogenicity)	NO
🔎	Structural Alert for nongenotoxic carcinogenicity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

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

OpenTox

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

OpenTox

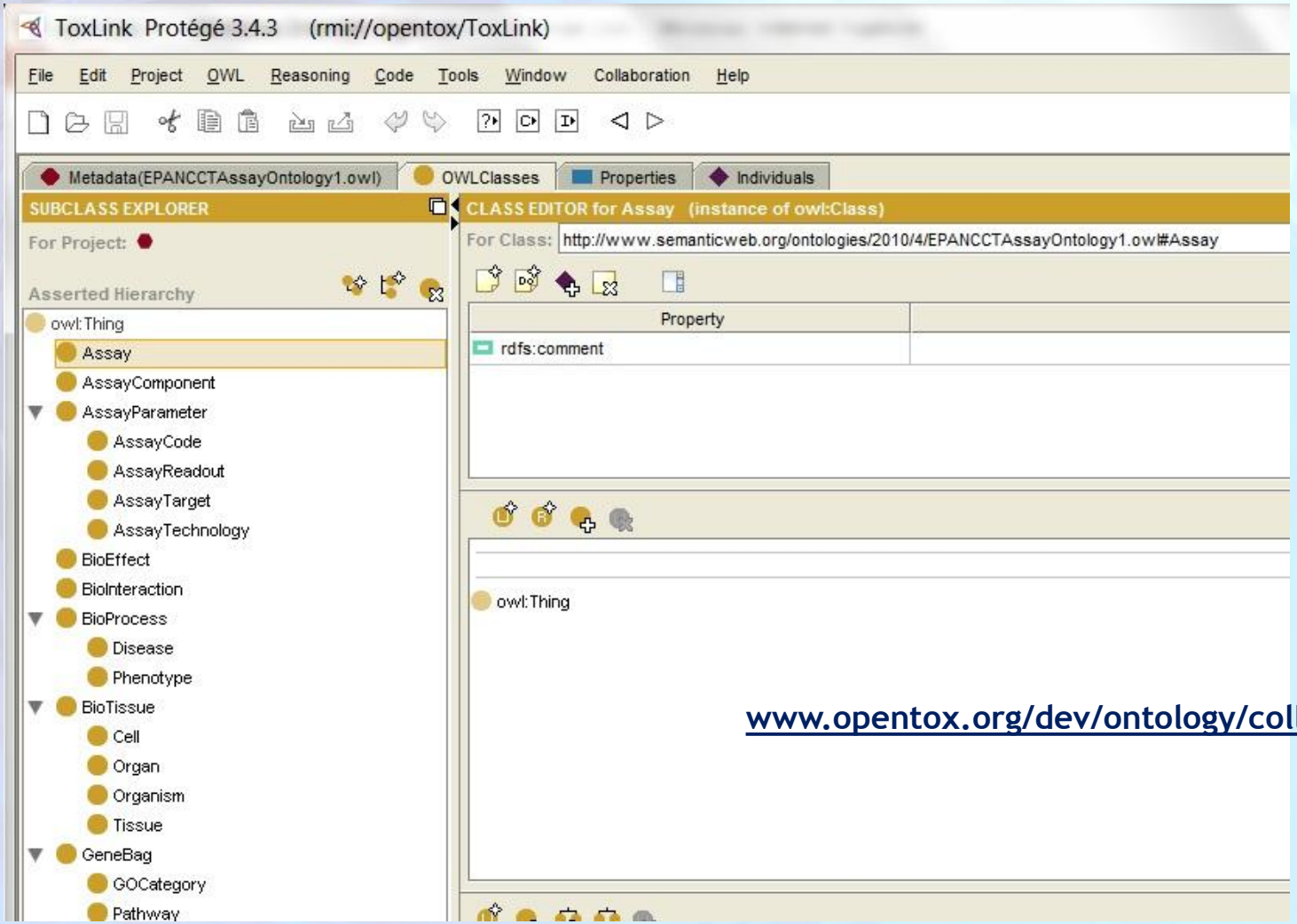
Information Requirements – *in vitro*

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

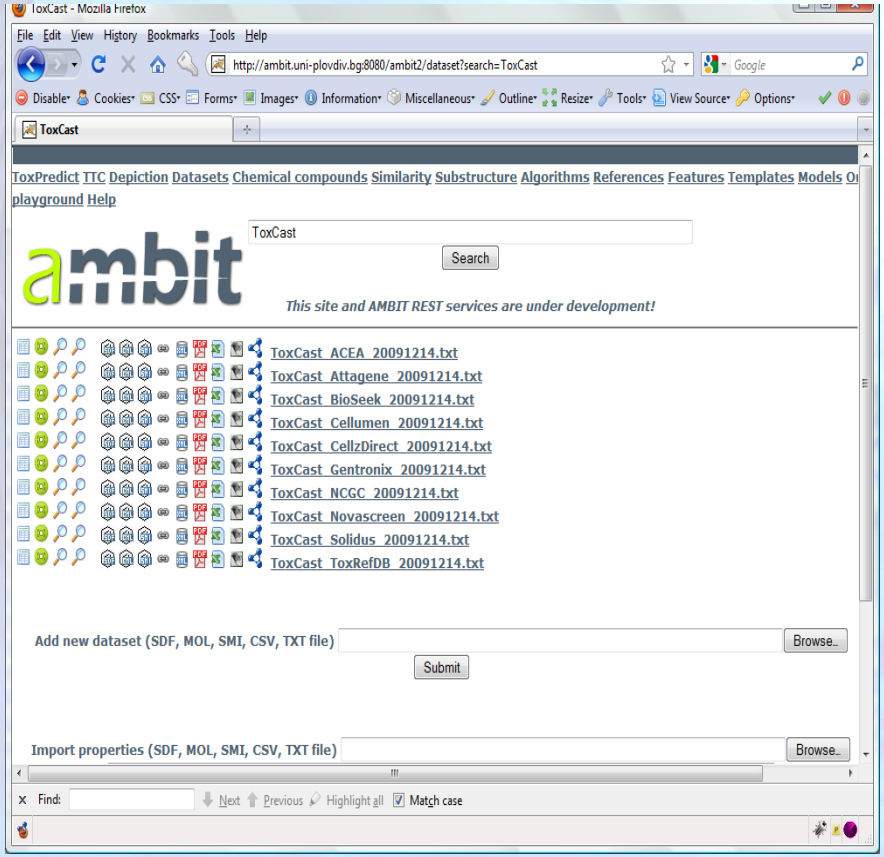


www.opentox.org/dev/ontology/collaborative_protege

OpenTox

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 rdf:type toxcast:SPECIES}.
  {?assay toxcast:hasProperty ?target_source}.
  {?target_source rdf:type toxcast:ASSAY_TARGET_SOURCE}
  {?assay toxcast:hasProperty ?target_family}.
  {?target_family rdf:type toxcast:ASSAY_TARGET_FAMILY}.
  {?assay toxcast:hasProperty ?target}.
  {?target rdf:type toxcast:ASSAY_TARGET}.
   {?assay toxcast:hasProperty toxcast:Cytochrome_P450}.
                }
order by ?feature ?assay ?target
```

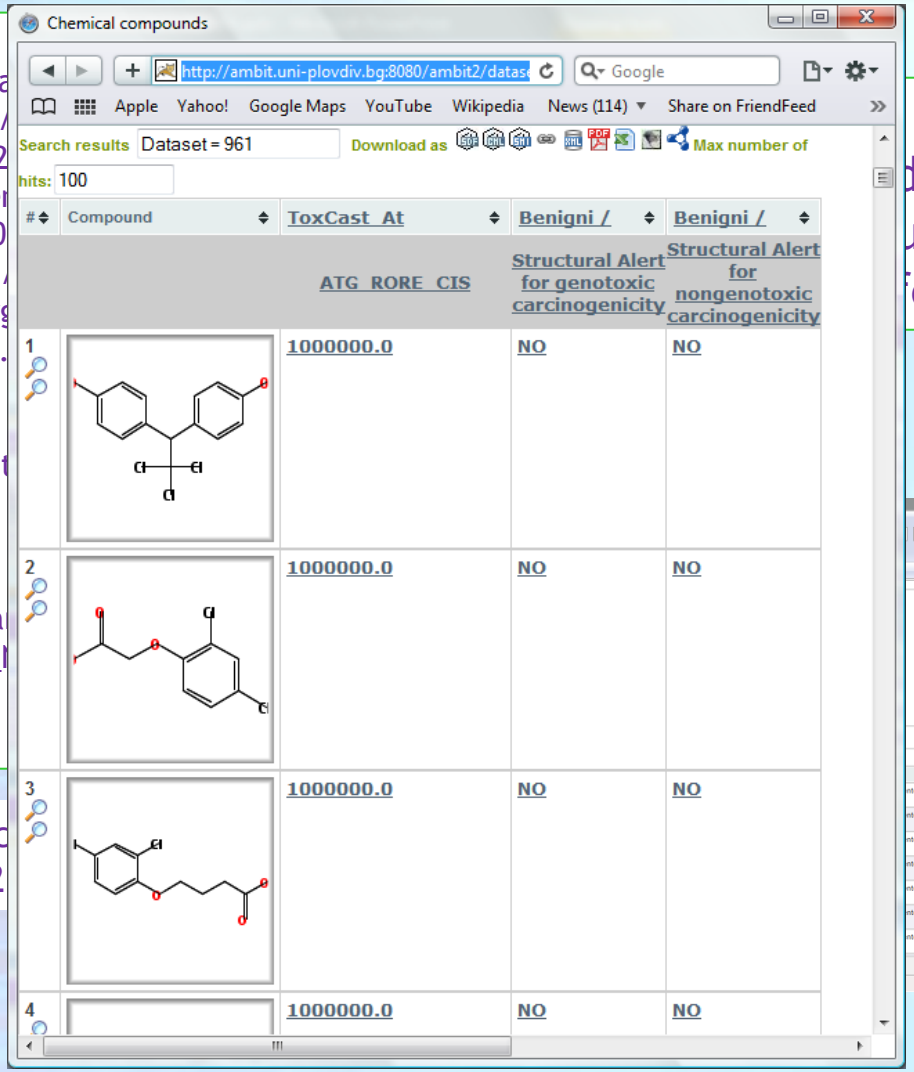
OpenTox

Example: ToxCast

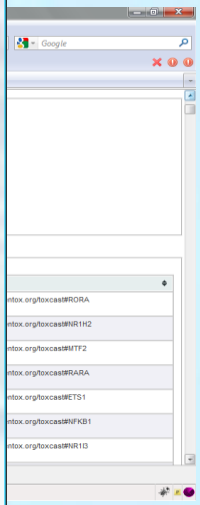
```
PREFIX ot:<http://www.opentox.org/a
PREFIX ota:<http://www.opentox.org/
PREFIX owl:<http://www.w3.org/2002
PREFIX dc:<http://purl.org/dc/elemer
PREFIX rdfs: <http://www.w3.org/200
PREFIX rdf:<http://www.w3.org/1999/
PREFIX otee:<http://www.opentox.org
PREFIX toxcast:<http://www.opentox.
                select ?Feature ?title ?id
                where {
                ?Feature rdf:type ot:Feat
  {?Feature dc:title ?title}.
  {?Feature owl:sameAs ?assay}.
  {?assay toxcast:gene ?geneid}.
  {?assay toxcast:hasProperty ?genena
  {?genename rdf:type toxcast:GENE_N

                }
```

Query an OpenTox ontology servic
http://ambit.uni-plovdiv.bg:8082

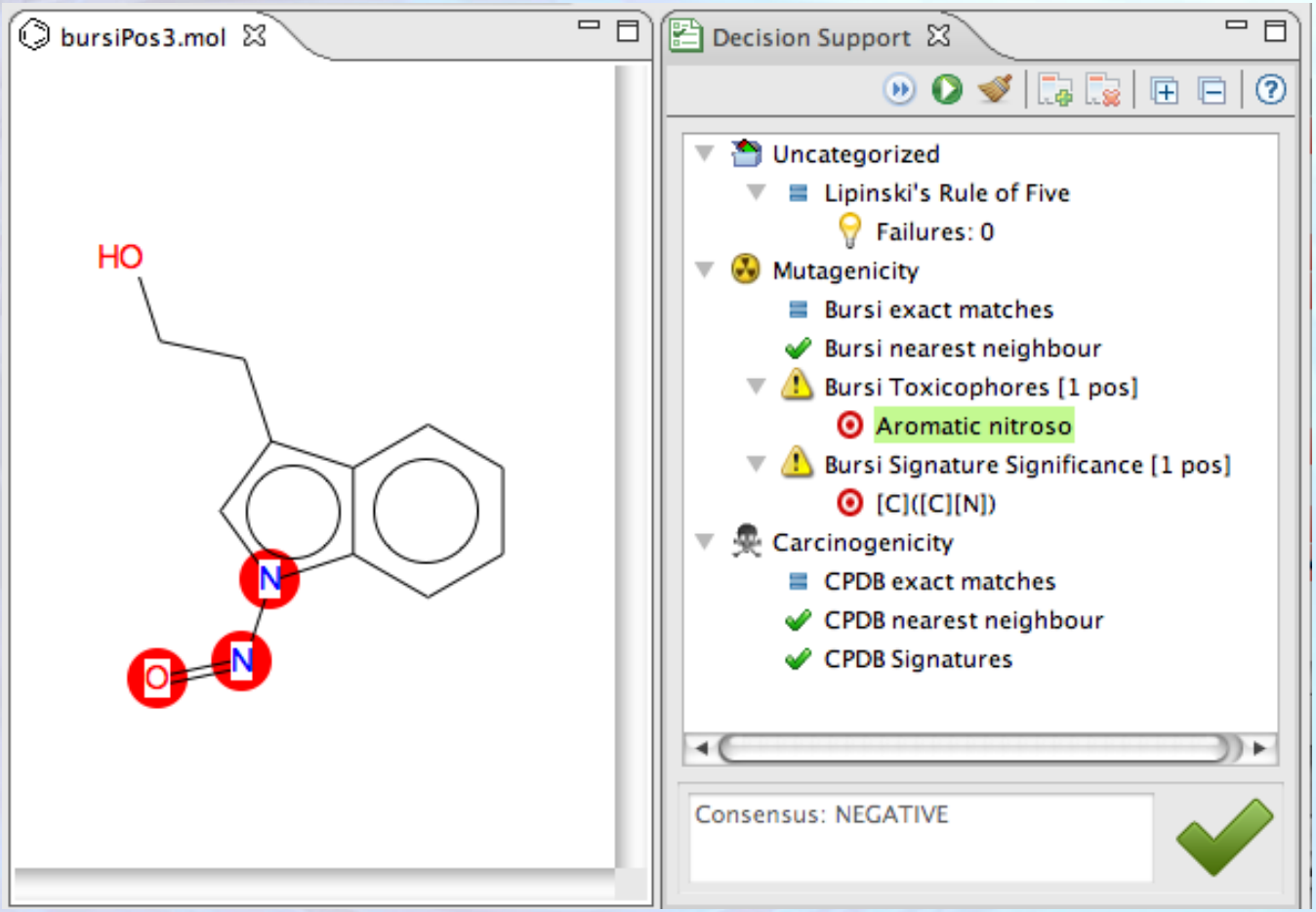


dataset/961?feat
uni-
feature/335126



OpenTox

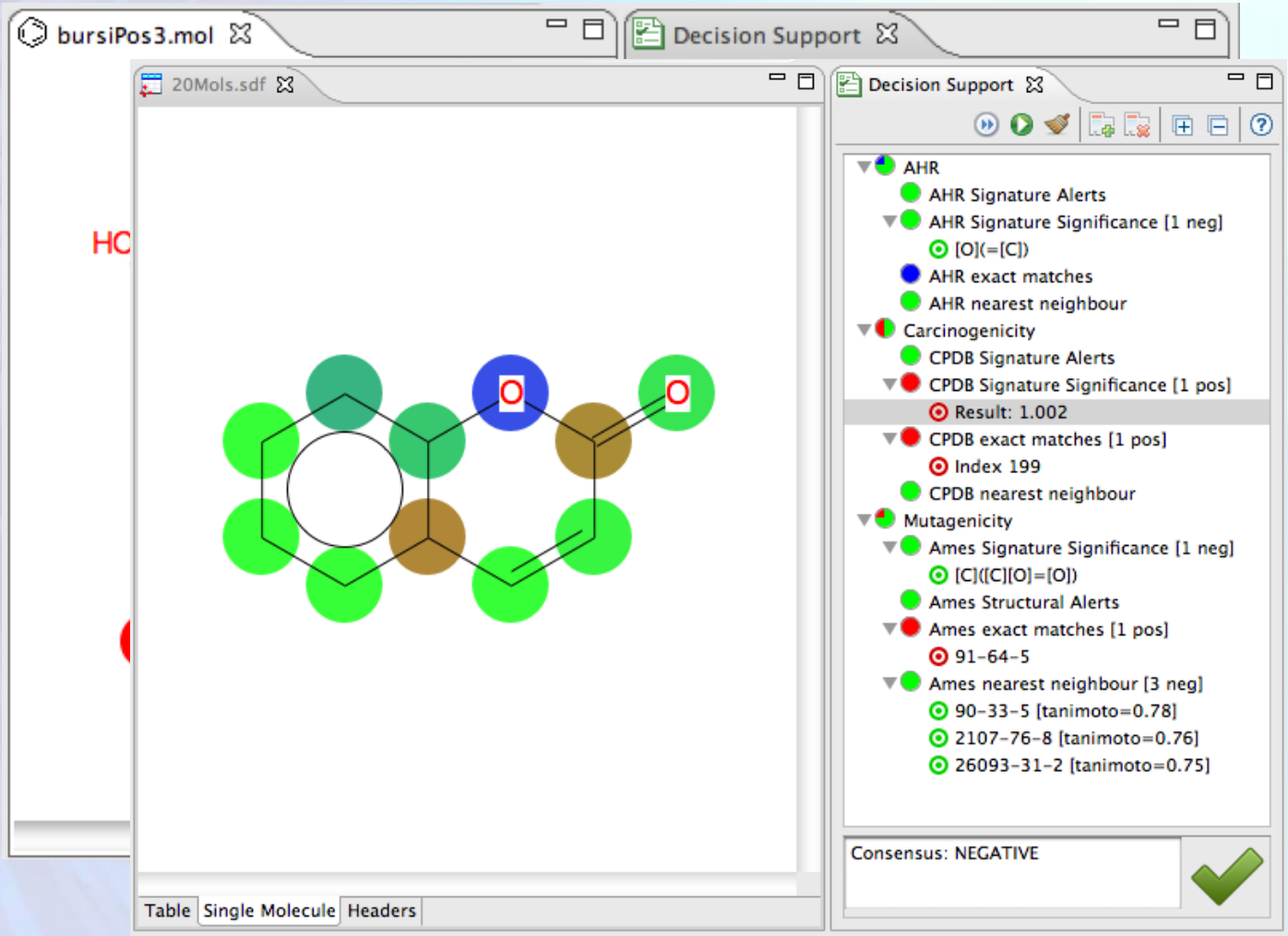
Bioclipse Visualisation Workbench



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

OpenTox

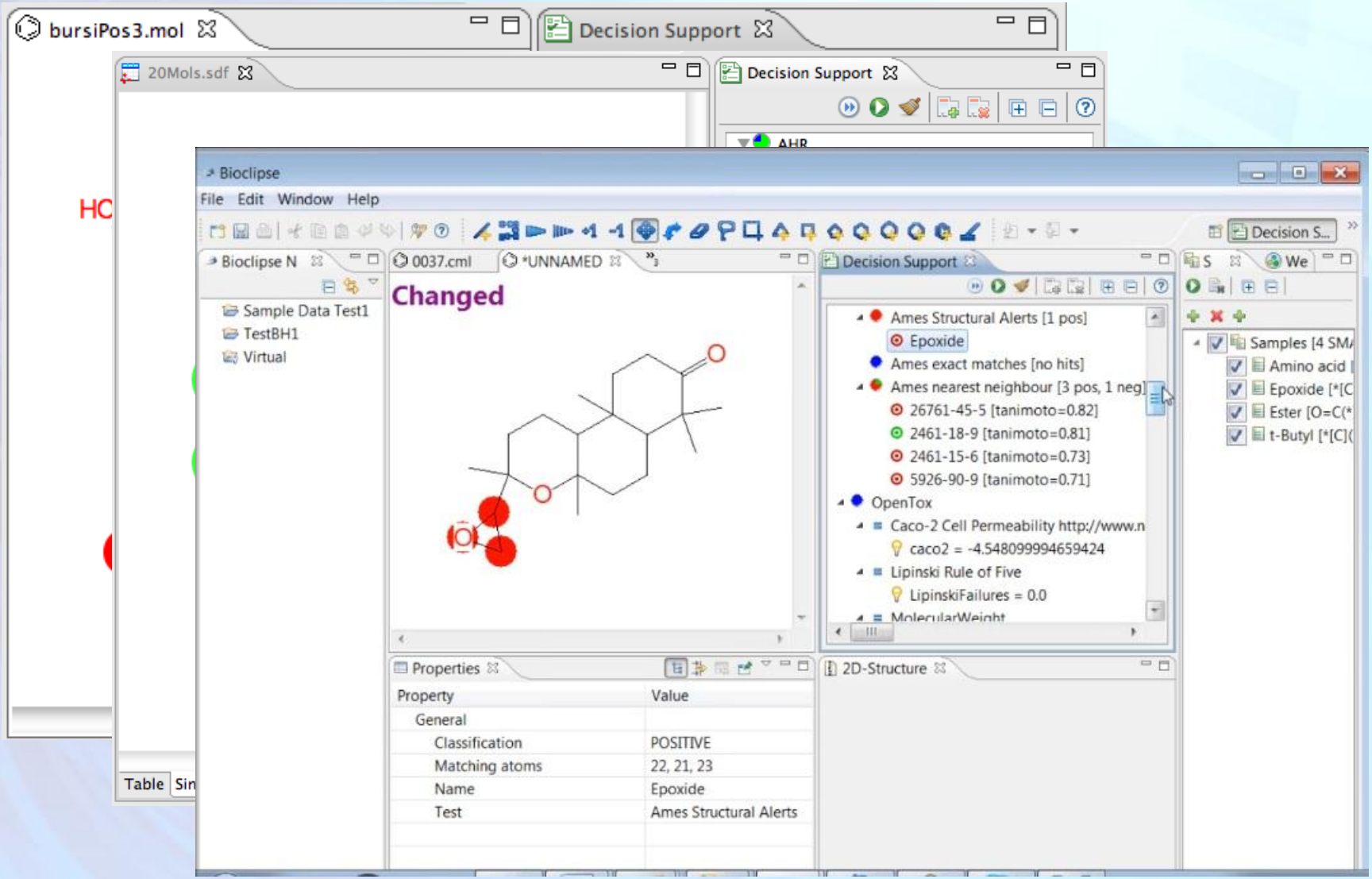
Bioclipse Visualisation Workbench



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

OpenTox

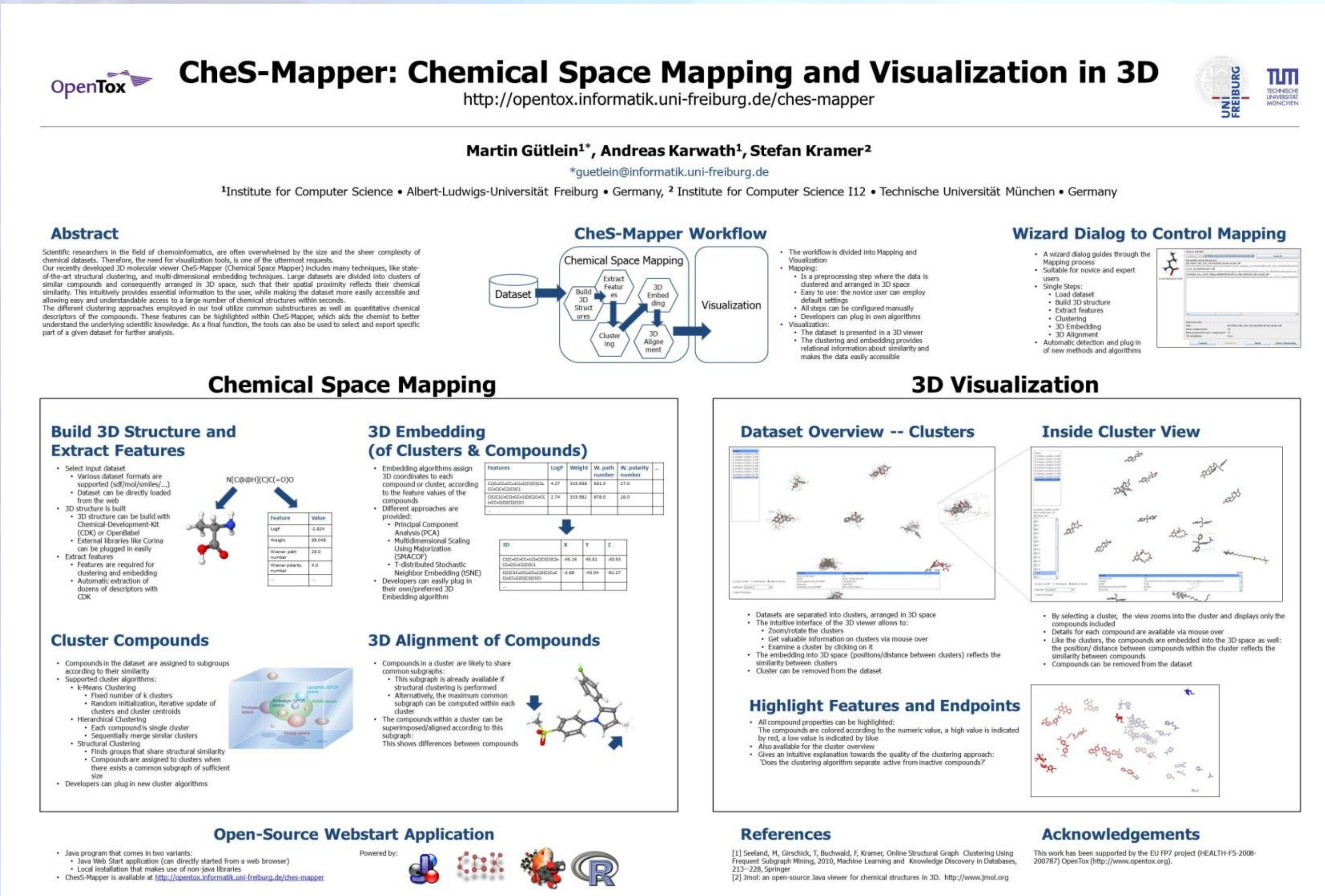
Bioclipse Visualisation Workbench - OpenTox



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

OpenTox

Chemical Space Visualisation (Ches-Mapper)



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

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

Martin Gütlein[1*], Andreas Karwath[1], Stefan Kramer[2]

*guetlein@informatik.uni-freiburg.de

[1]Institute for Computer Science • Albert-Ludwigs-Universität Freiburg • Germany, [2] Institute for Computer Science I12 • Technische Universität München • Germany

Abstract

Scientific researchers in the field of chemoinformatics, are often overwhelmed by the size and the sheer complexity of chemical datasets. Therefore, the need for visualization tools, is one of the uttermost 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.

CheS-Mapper Workflow

- The workflow is divided into Mapping and Visualization
- Mapping:
  - Is a preprocessing step where the data is clustered and arranged in 3D space
  - Easy to use: the novice user can employ default settings
  - All steps can be configured manually
  - Developers can plug in own algorithms
- Visualization:
  - The dataset is presented in a 3D viewer
  - The clustering and embedding provides relational information about similarity and makes the data easily accessible

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 (sdf/mol/smiles/...)
  - Dataset can be directly loaded from the web
- 3D structure is built
  - 3D structure can be build 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

N[C@@H](C)C(=O)O

Feature	Value
LogP	-2.824
Weight	89.048
Wiener path number	29.0
Wiener polarity number	4.0
...	...

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)
  - Multidimensional Scaling Using Majorization (SMACOF)
  - T-distributed Stochastic Neighbor Embedding (tSNE)
- Developers can easily plug in their own/preferred 3D Embedding algorithm

Features	LogP	Weight	W. path number	W. polarity number	...
[illegible]	4.27	315.936	581.0	27.0	
[illegible]	2.74	315.981	678.0	28.0	
...					

3D	X	Y	Z
[illegible]	-45.16	45.81	-30.53
[illegible]	-2.68	-43.04	-81.27
...			

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 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
- ChesS-Mapper is available at http://opentox.informatik.uni-freiburg.de/ches-mapper

Powered by:

References

[1] Seeland, M, Girschick, T, Buchwald, F, Kramer, Online Structural Graph Clustering Using Frequent Subgraph Mining, 2010, Machine Learning and Knowledge Discovery in Databases, 213--228, Springer
[2] Jmol: an open-source Java viewer for chemical structures in 3D. http://www.jmol.org

Acknowledgements

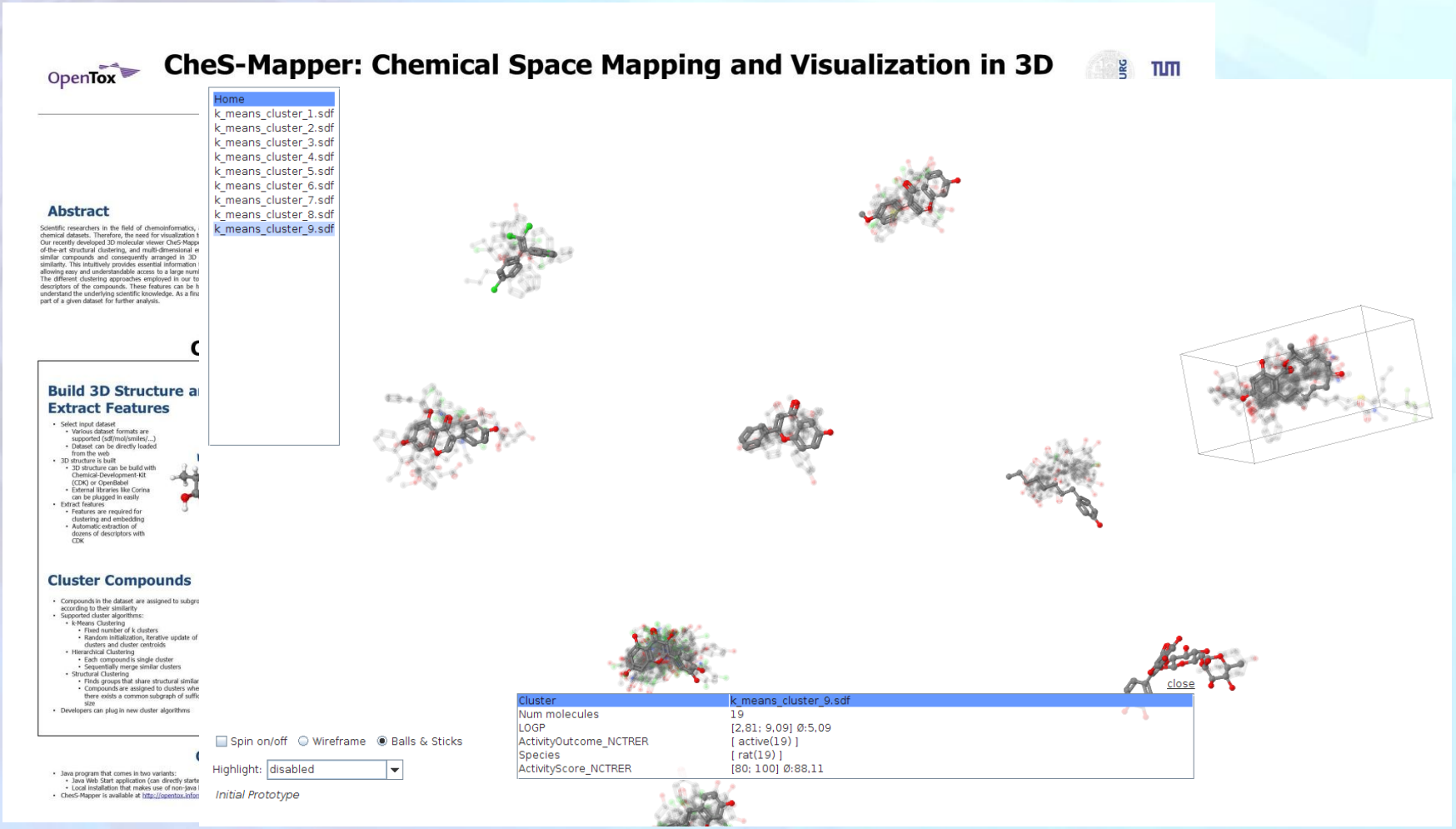
This work has been supported by the EU FP7 project (HEALTH-F5-2008-200787) OpenTox (http://www.opentox.org).

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

OpenTox

SEVENTH FRAMEWORK PROGRAMME

Chemical Space Visualisation (Ches-Mapper)



OpenTox

Forming chemical feature-based categories

OpenTox

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

Abstract

Scientific researchers in the field of chemoinfom
chemical datasets. Therefore, the need for visuali
Our recently developed 3D molecular viewer Ches
of-the-art structural clustering, and multi-dimens
similar compounds and consequently arranged
similarity. This intuitively provides essential infor
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.

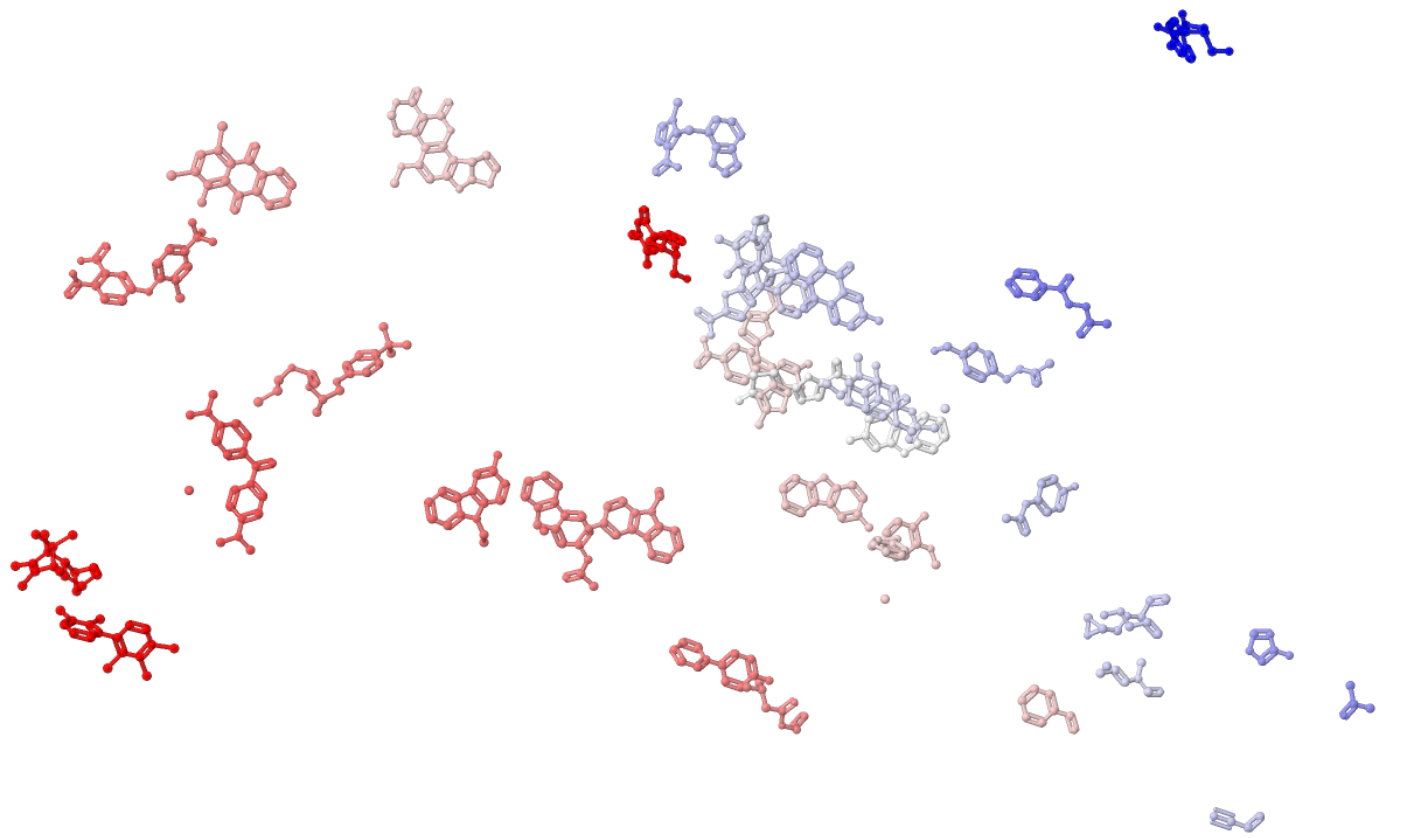
Build 3D Structure
Extract Features

- Select input dataset
  - Various dataset formats are supported (sdf/mol/smiles/...)
  - Dataset can be directly loaded from the web
- 3D structure is built
  - 3D structure can be build 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 according 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 cluste there exists a common subgraph size
- Developers can plug in new cluster algorit

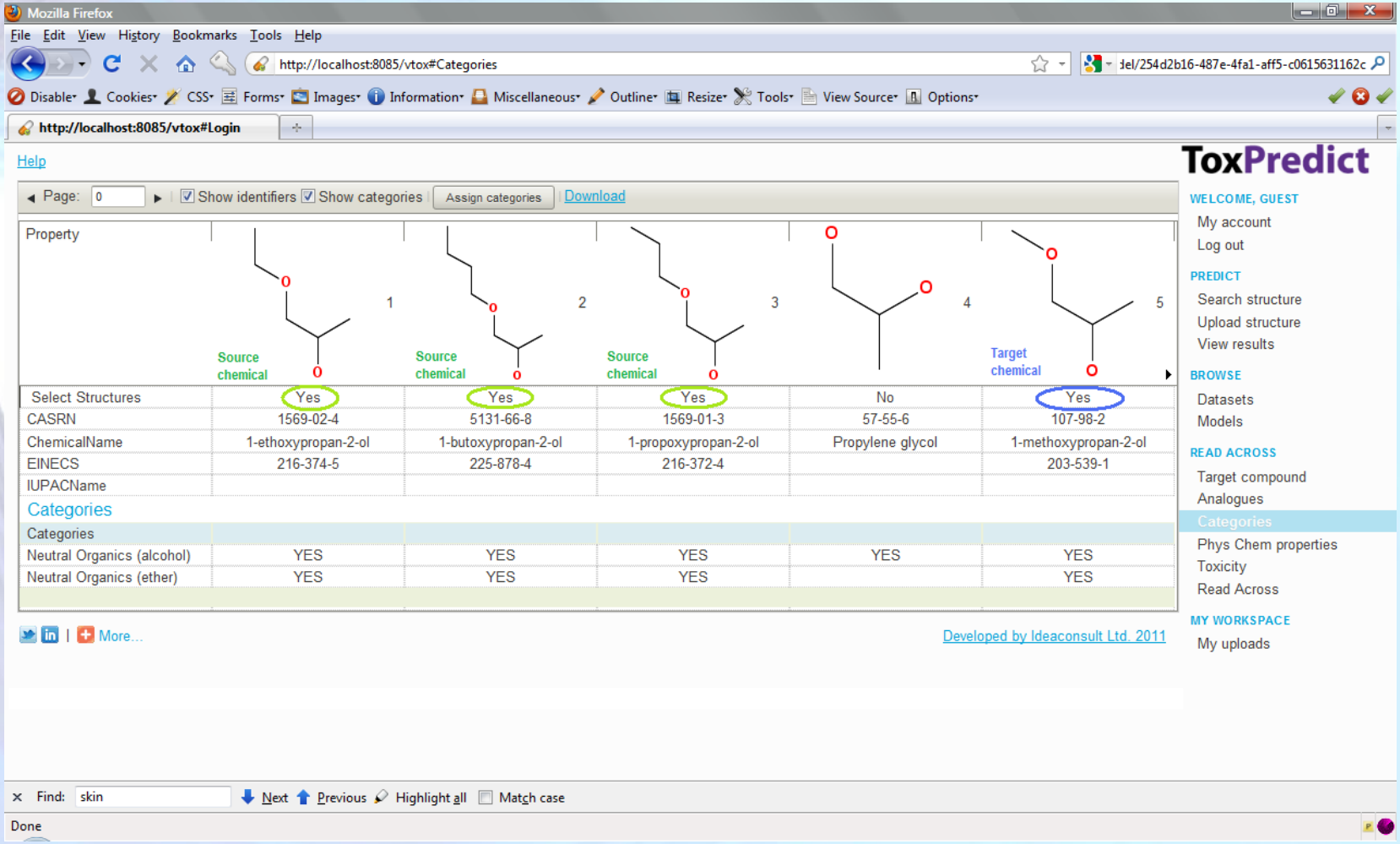
- Java program that comes in two variants:
  - Java Web Start application (can direct
  - Local installation that makes use of no
- CheS-Mapper is available at http://opente



OpenTox

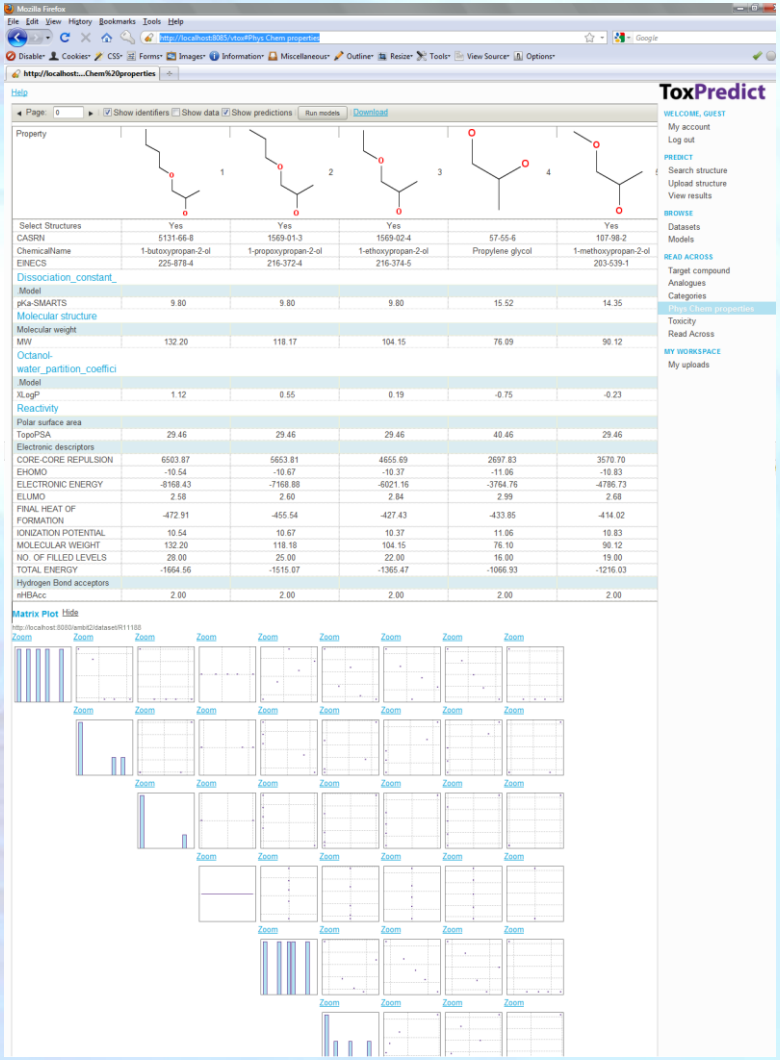
SEVENTH FRAMEWORK PROGRAMME

Read-Across Application



OpenTox

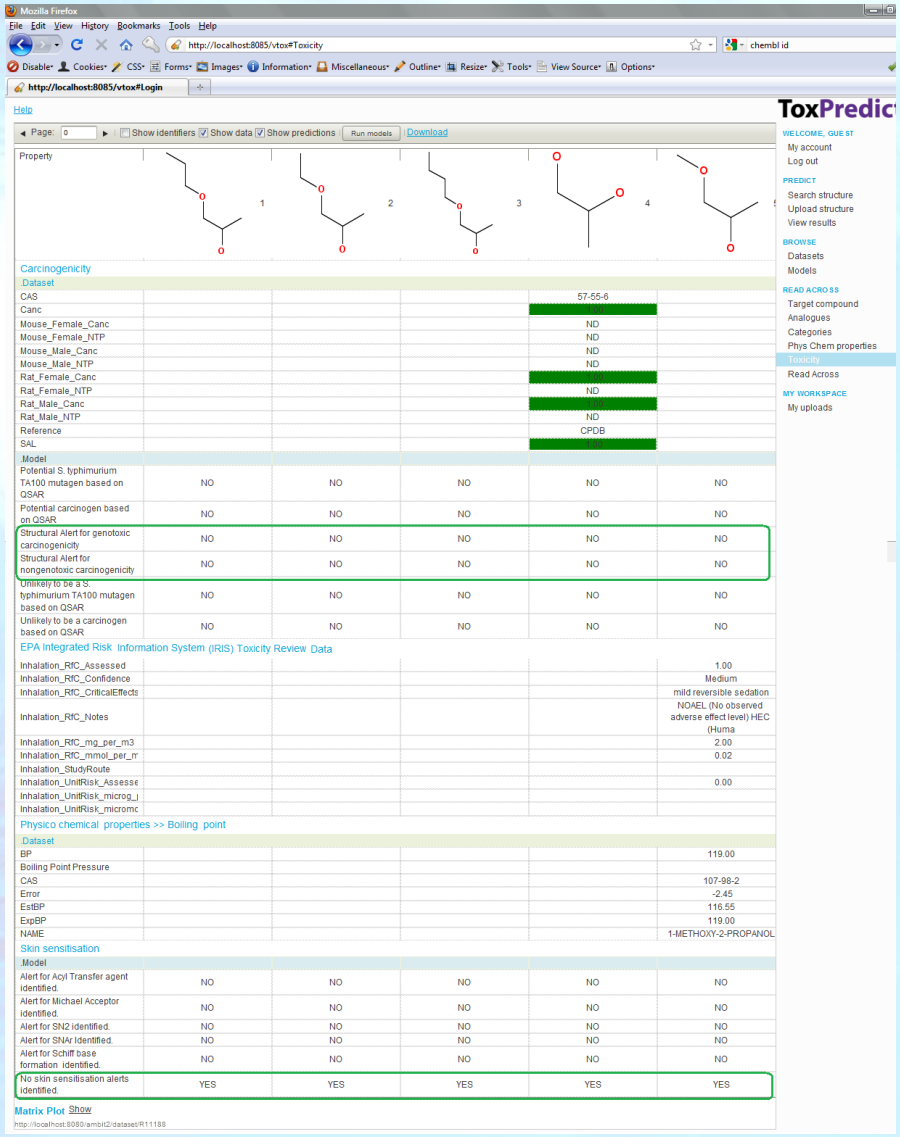
Read-Across Application



OpenTox

SEVENTH FRAMEWORK PROGRAMME

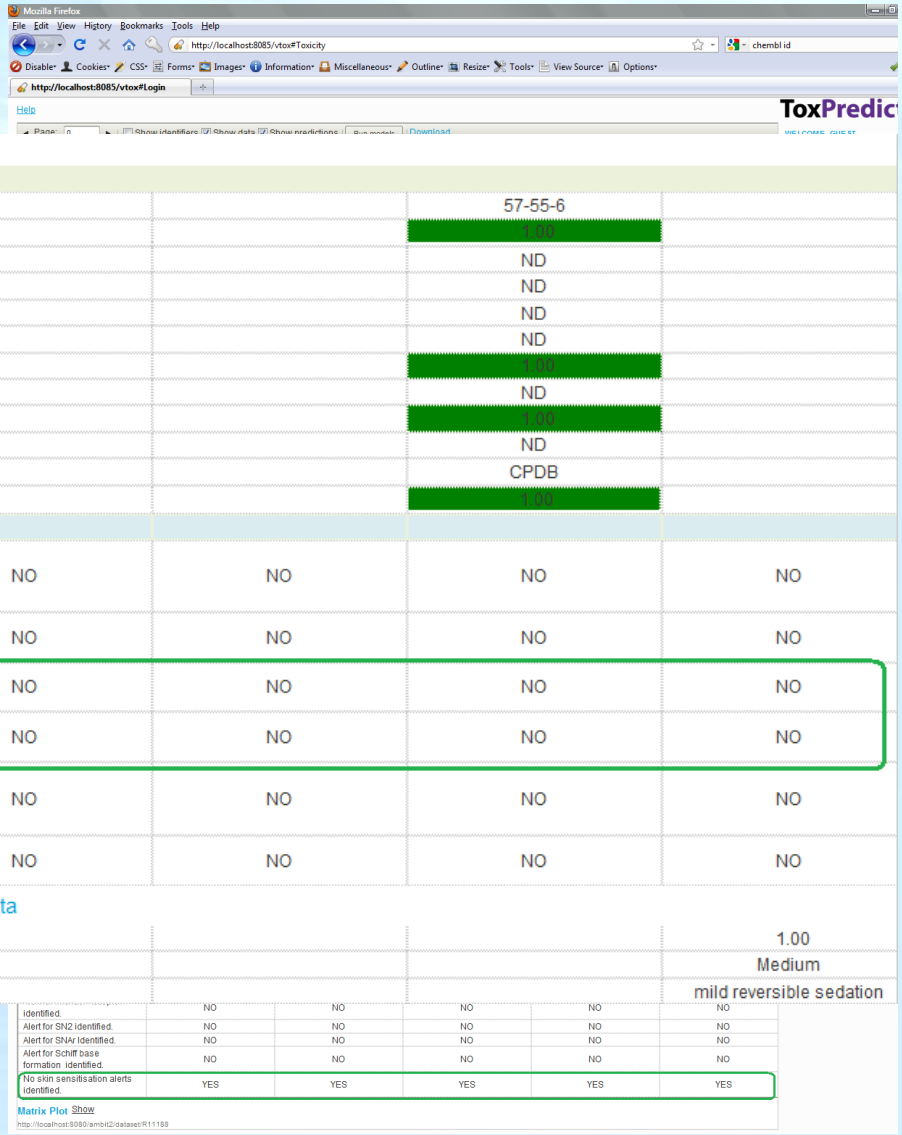
Read-Across Application



OpenTox

SEVENTH FRAMEWORK PROGRAMME

Read-Across Application



Carcinogenicity

.Dataset					
CAS				57-55-6	
Canc				1.00	
Mouse_Female_Canc				ND	
Mouse_Female_NTP				ND	
Mouse_Male_Canc				ND	
Mouse_Male_NTP				ND	
Rat_Female_Canc				1.00	
Rat_Female_NTP				ND	
Rat_Male_Canc				1.00	
Rat_Male_NTP				ND	
Reference				CPDB	
SAL				1.00	
.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 SNAr 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

Matrix Plot Show

OpenTox

Ontology - automating consistency

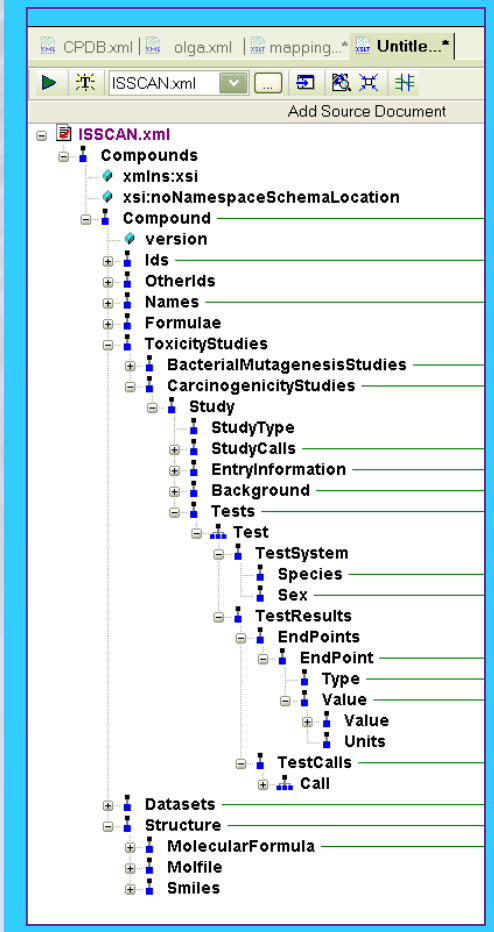
We have defined an ontology based on the OWL (Web Ontology Language) for toxicological endpoints which is in line with current ECHA REACH guidance.

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

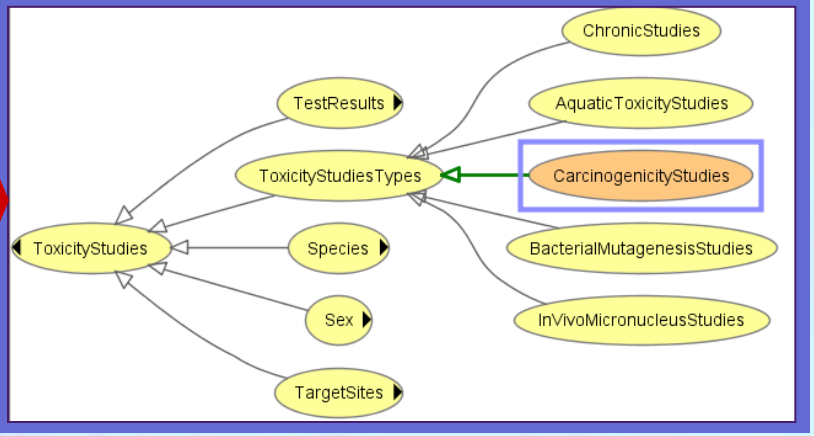
ToxML schema



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

OpenTox Toxicological Endpoint Ontology

Ontology Development



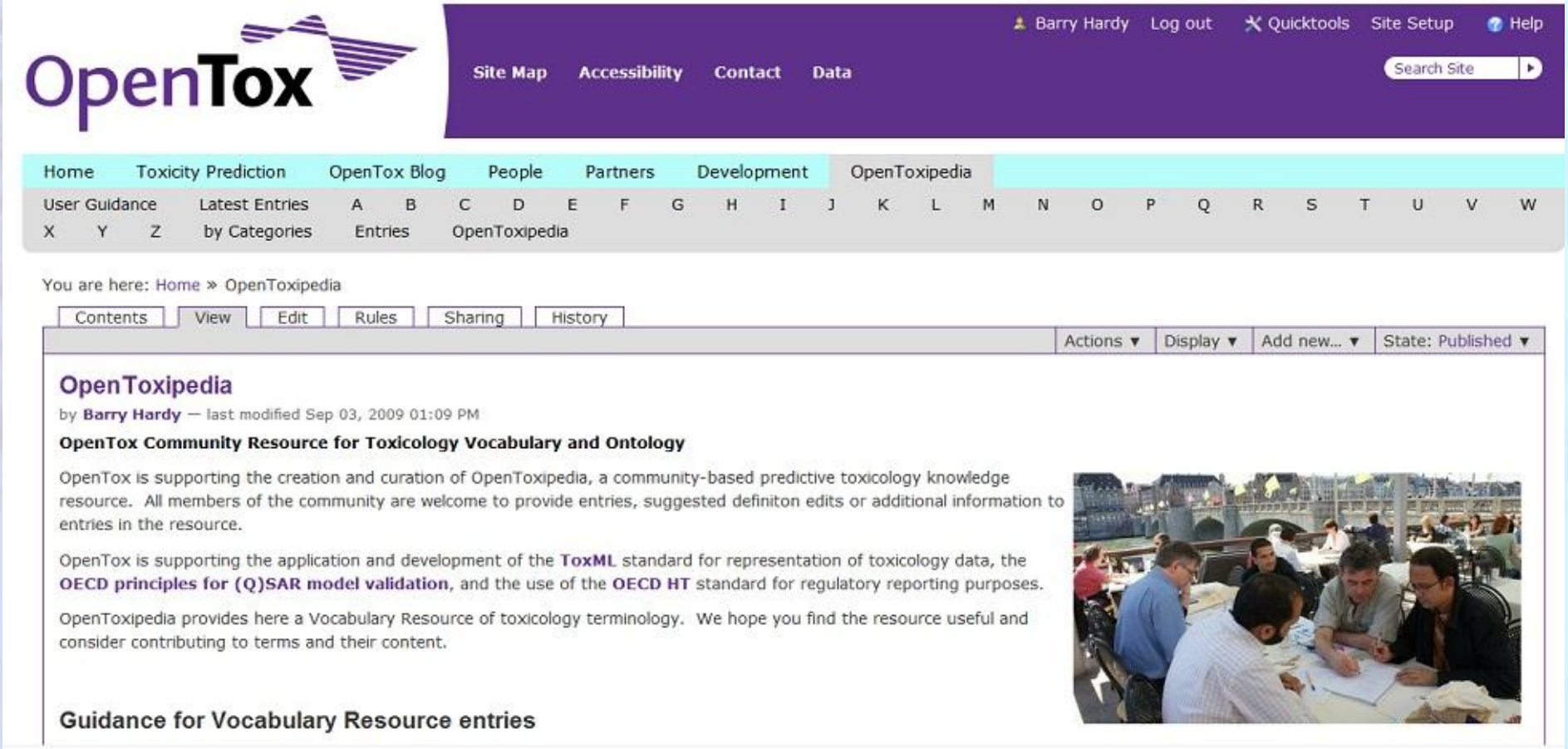
Collaborative Protégé Environment

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

www.opentox.org/dev/ontology/collaborative_protege

OpenTox

OpenToxipedia



www.opentox.org/opentoxipedia

OpenTox

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

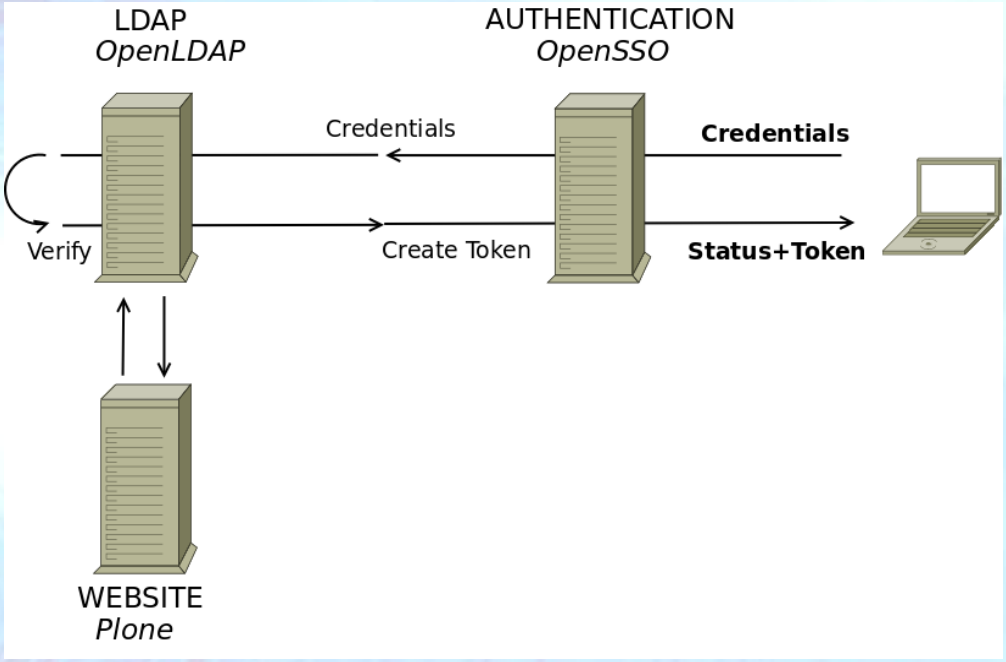
16 -17th November 2010, Hinxton, UK

OpenTox

Controlling Access to Confidential Information

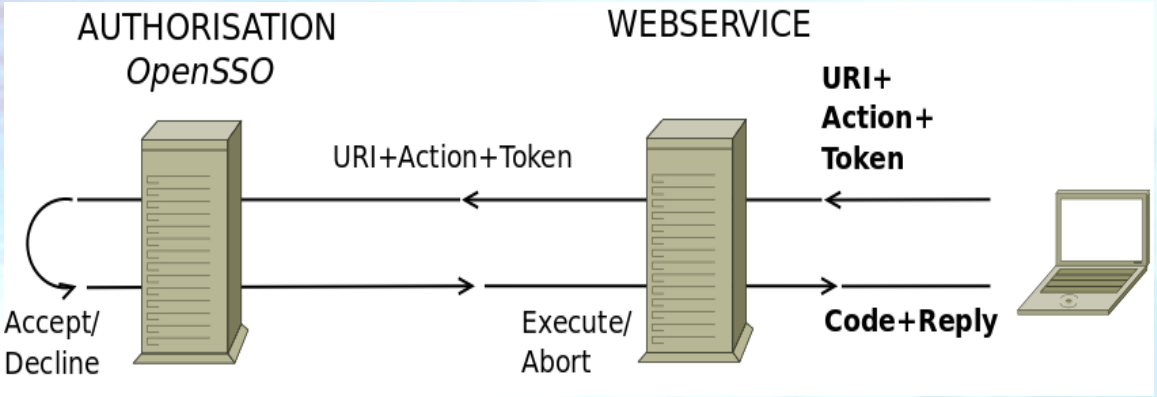
- OpenTox makes resources available through URIs
- OpenTox provides facilities to protect confidential information located at URIs. Two tasks are involved here:
  - Authentication: Confirming the identity of the user requesting access
  - Authorisation: Granting the confirmed identity access according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

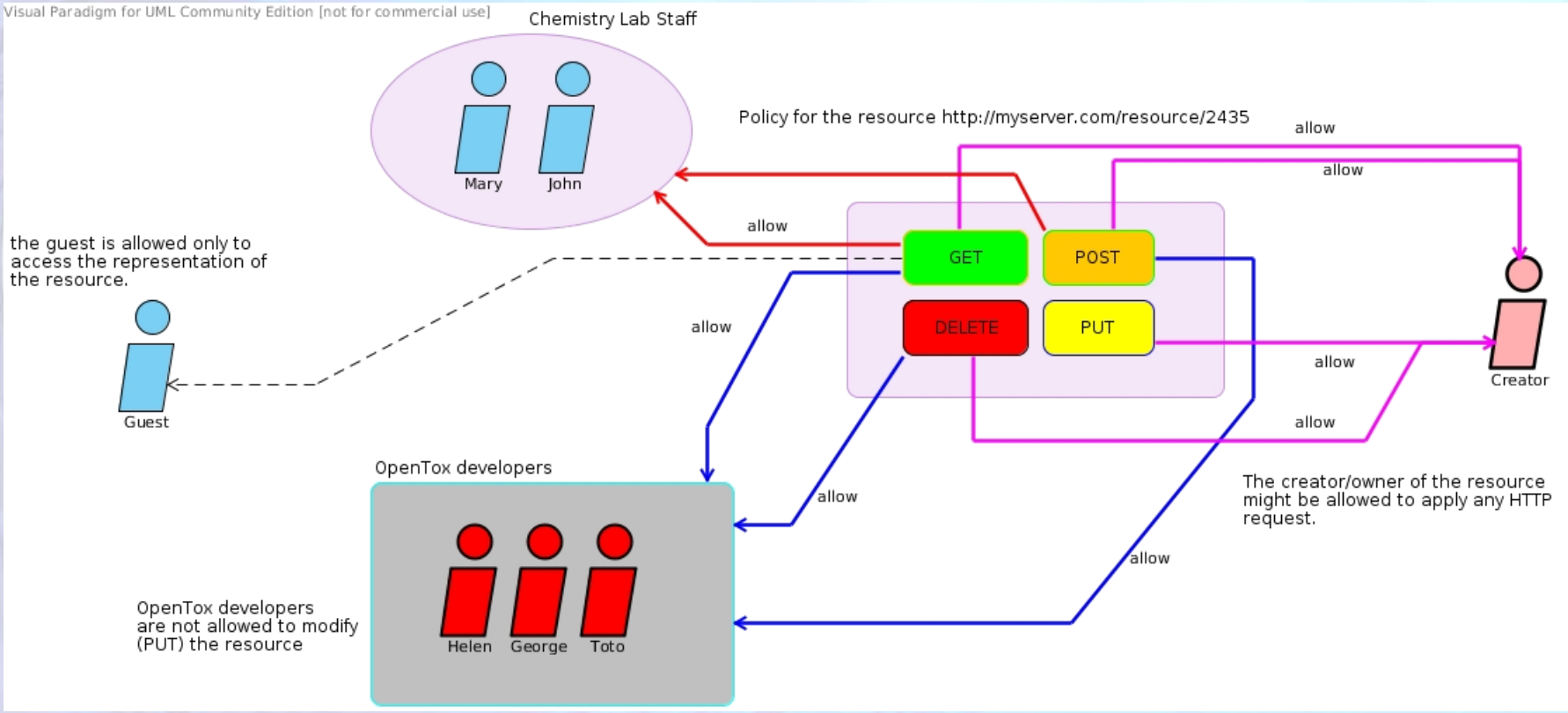
Authorisation



- 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

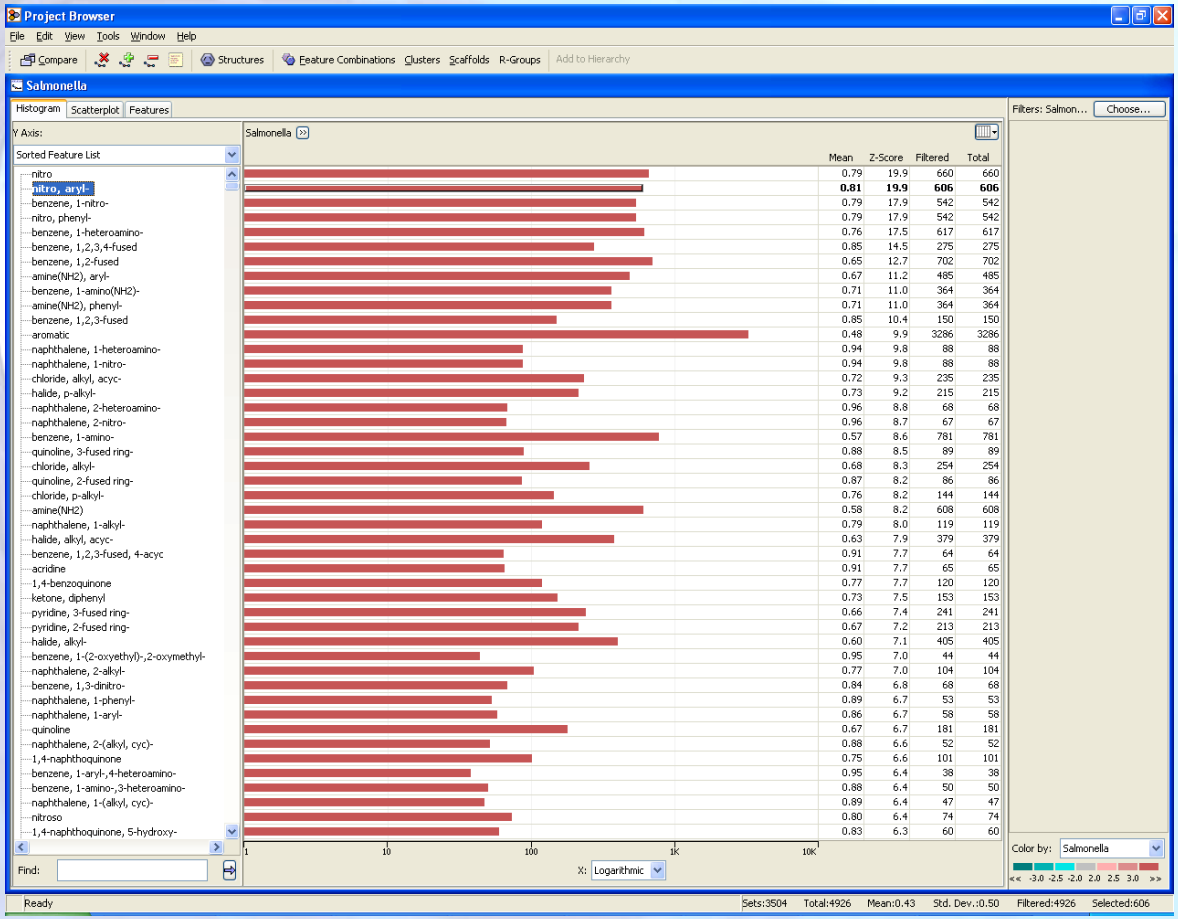
OpenTox

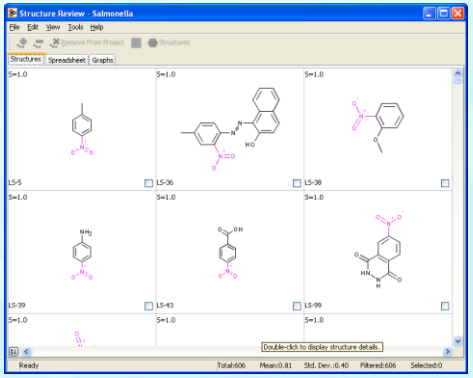
Policies

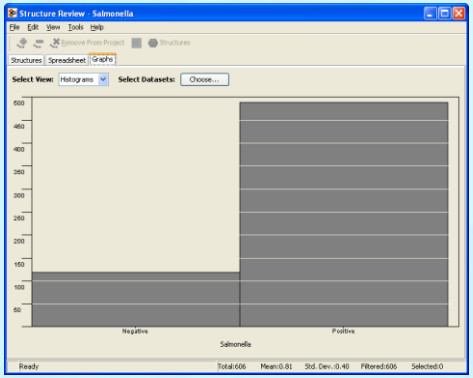


OpenTox

OpenTox - Leadscope Integration



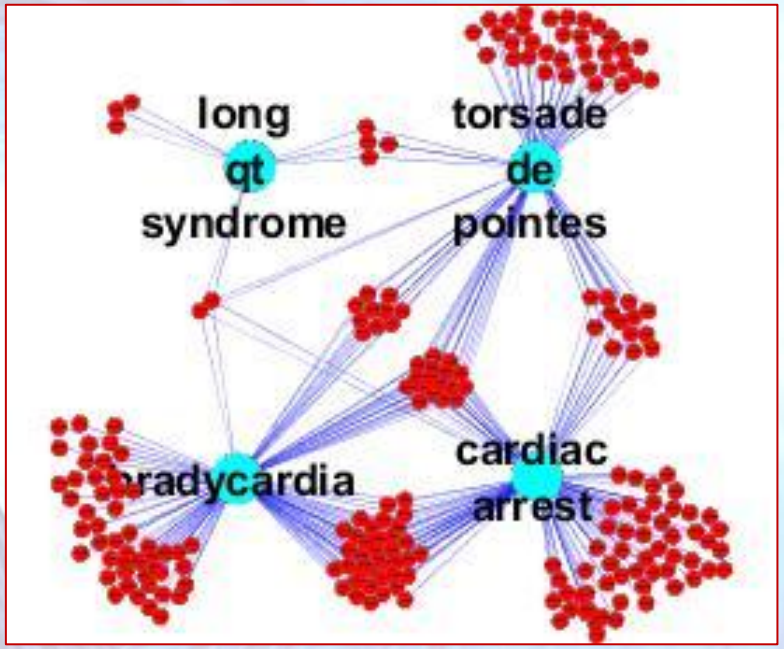




OpenTox

SEVENTH FRAMEWORK PROGRAMME

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?

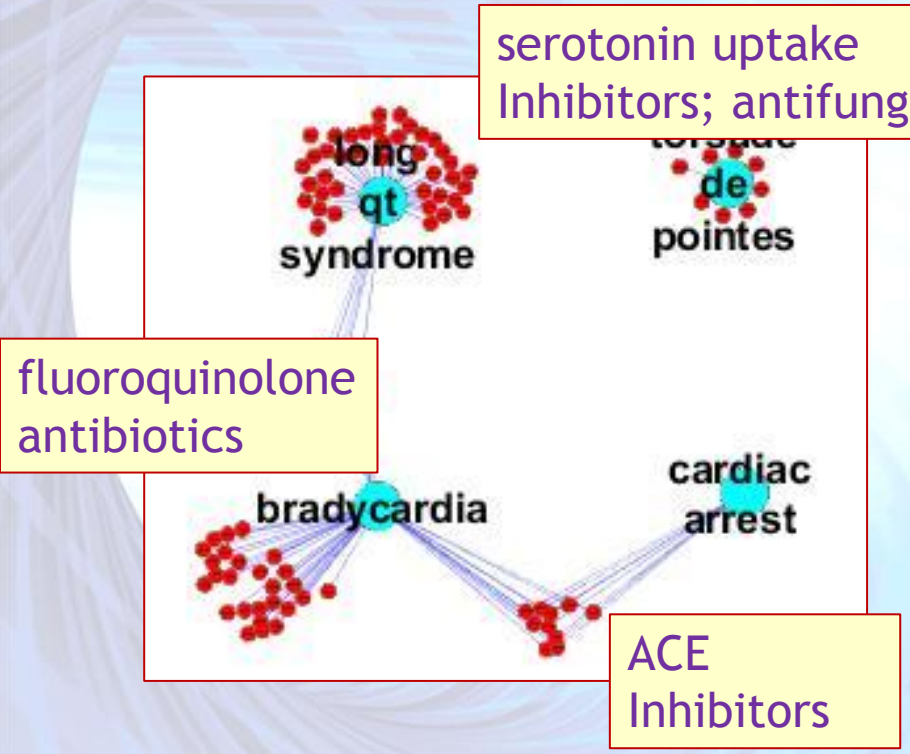
- Cardiac adverse events
- Related to hERG ion channel?

cyan = adverse event, red = drug
lines define links

PHARMATROPE

OpenTox

Example: Cardiac Adverse Events



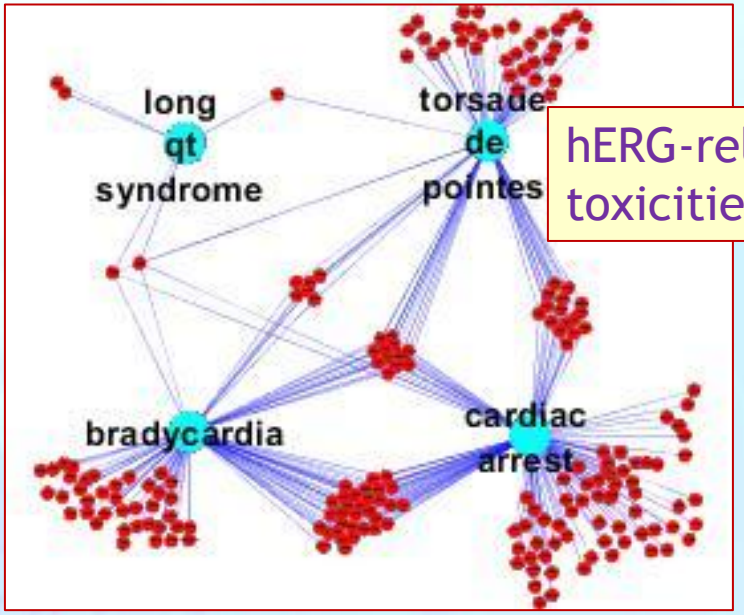
serotonin uptake Inhibitors; antifungals

fluoroquinolone antibiotics

ACE Inhibitors

On-target events

cyan = adverse event, red = drug lines define links



hERG-related toxicities?

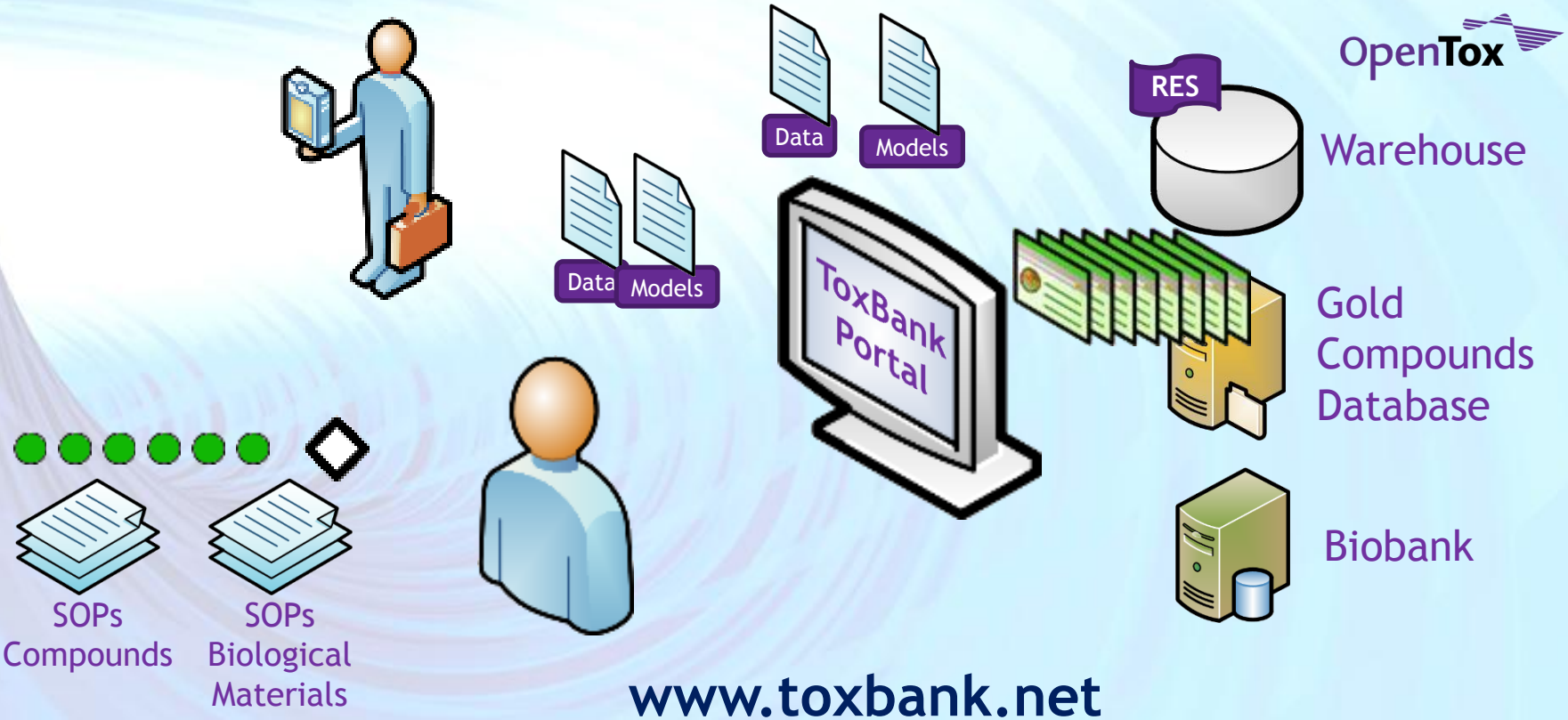
Off-target events

PHARMATROPE

OpenTox

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



www.toxbank.net

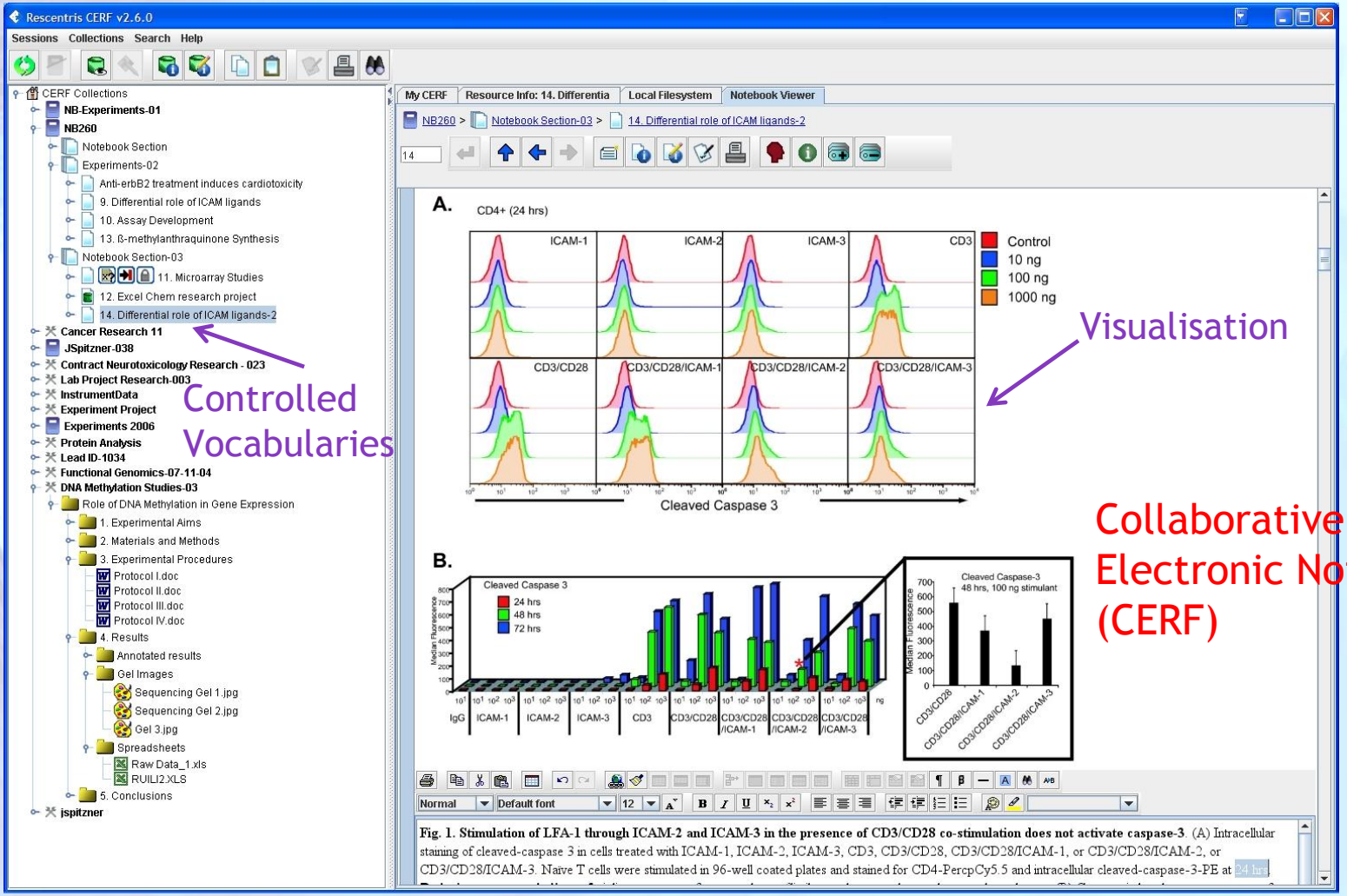
OpenTox

This project is jointly funded by COLIPA and the EC. Any opinions expressed in this slide are those of the author. COLIPA is not liable for any use that may be made of the information contained therein.

COLIPA The European Cosmetics Association

SEVENTH FRAMEWORK PROGRAMME

Collaborative Research Framework Integration



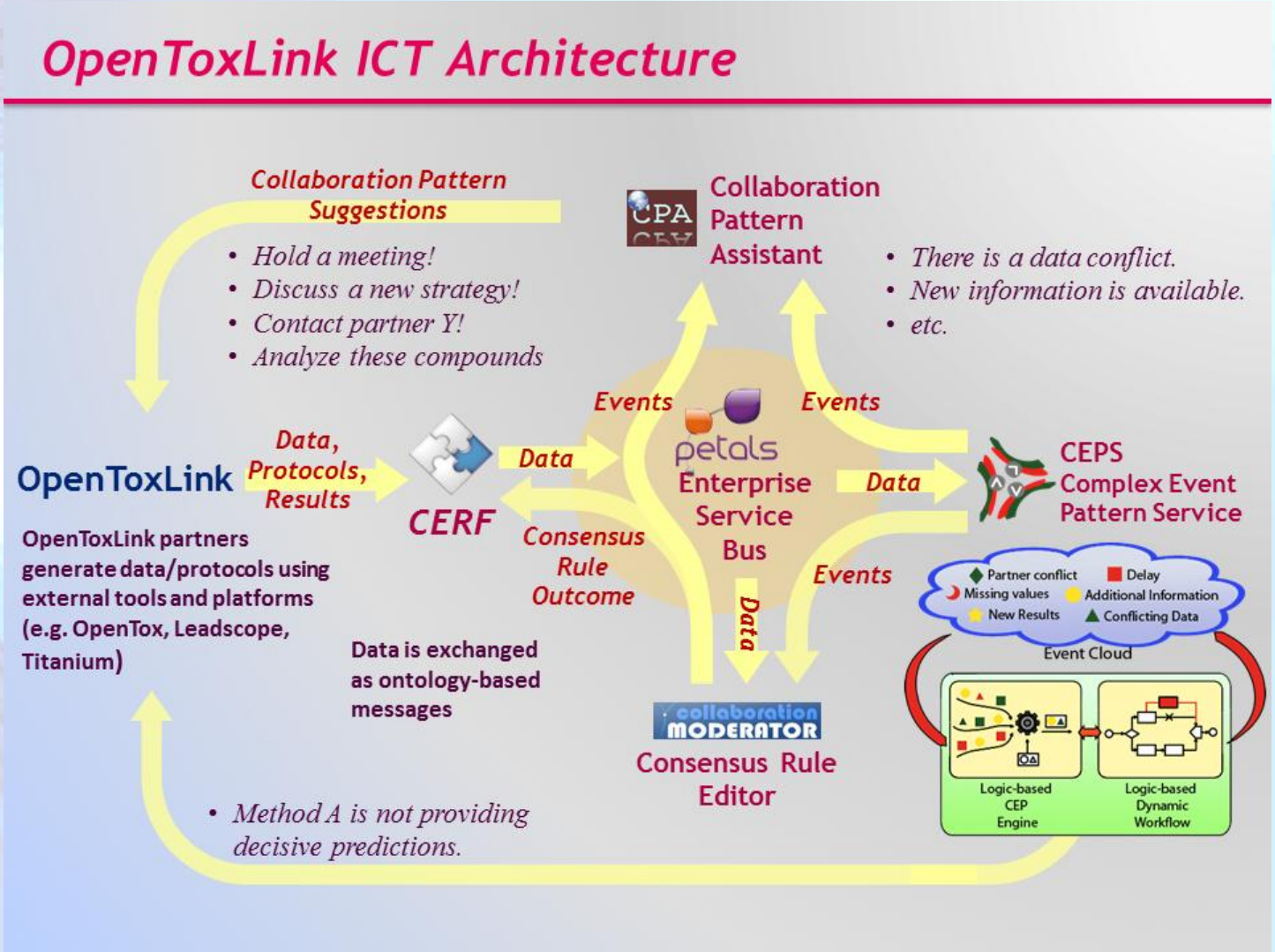
Controlled Vocabularies

Visualisation

Collaborative Electronic Notebook (CERF)

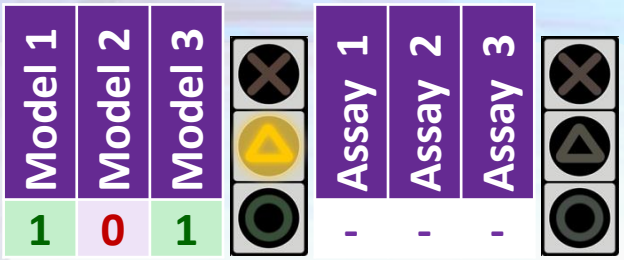
OpenTox

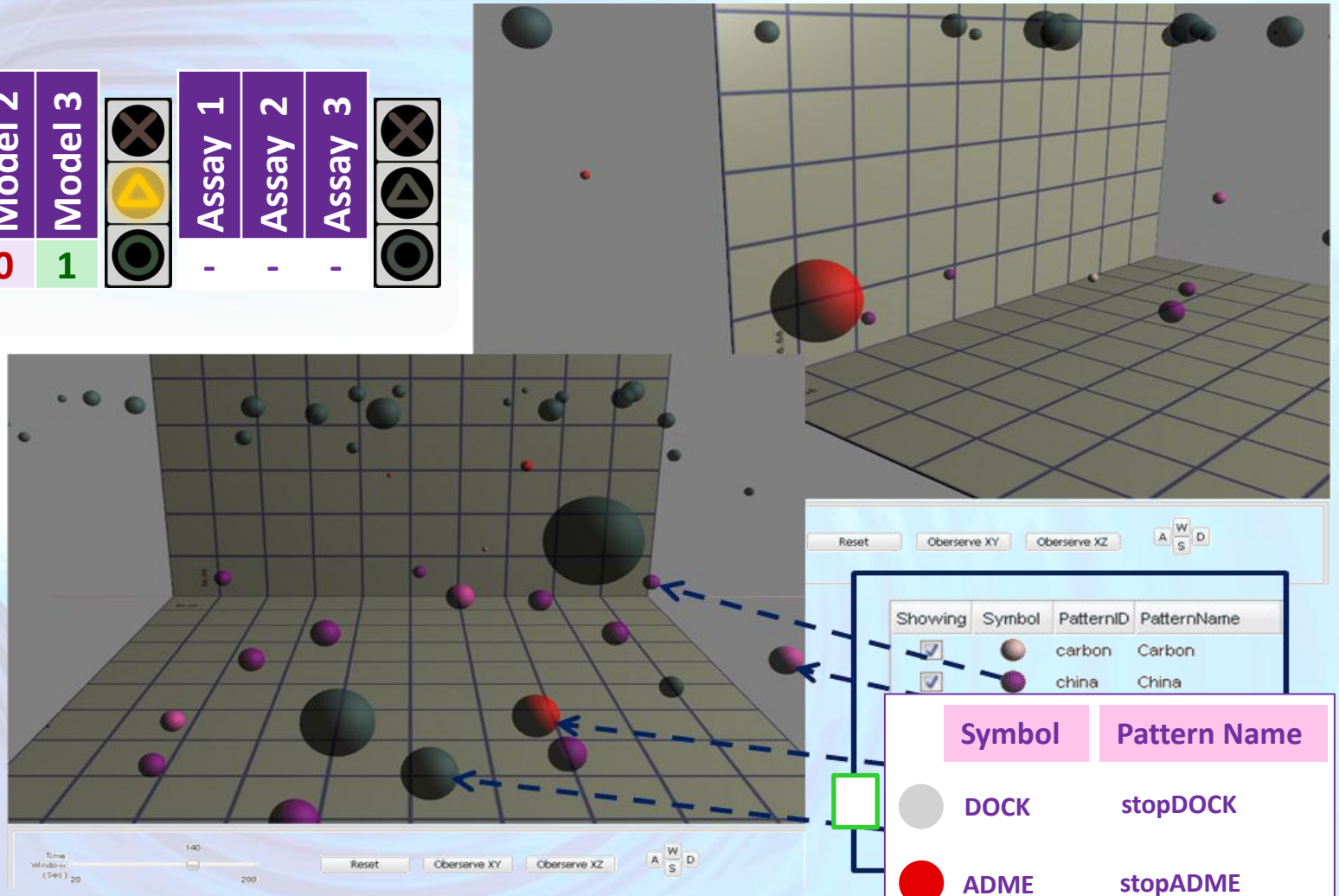
Event Driven Collaboration Architecture



OpenTox

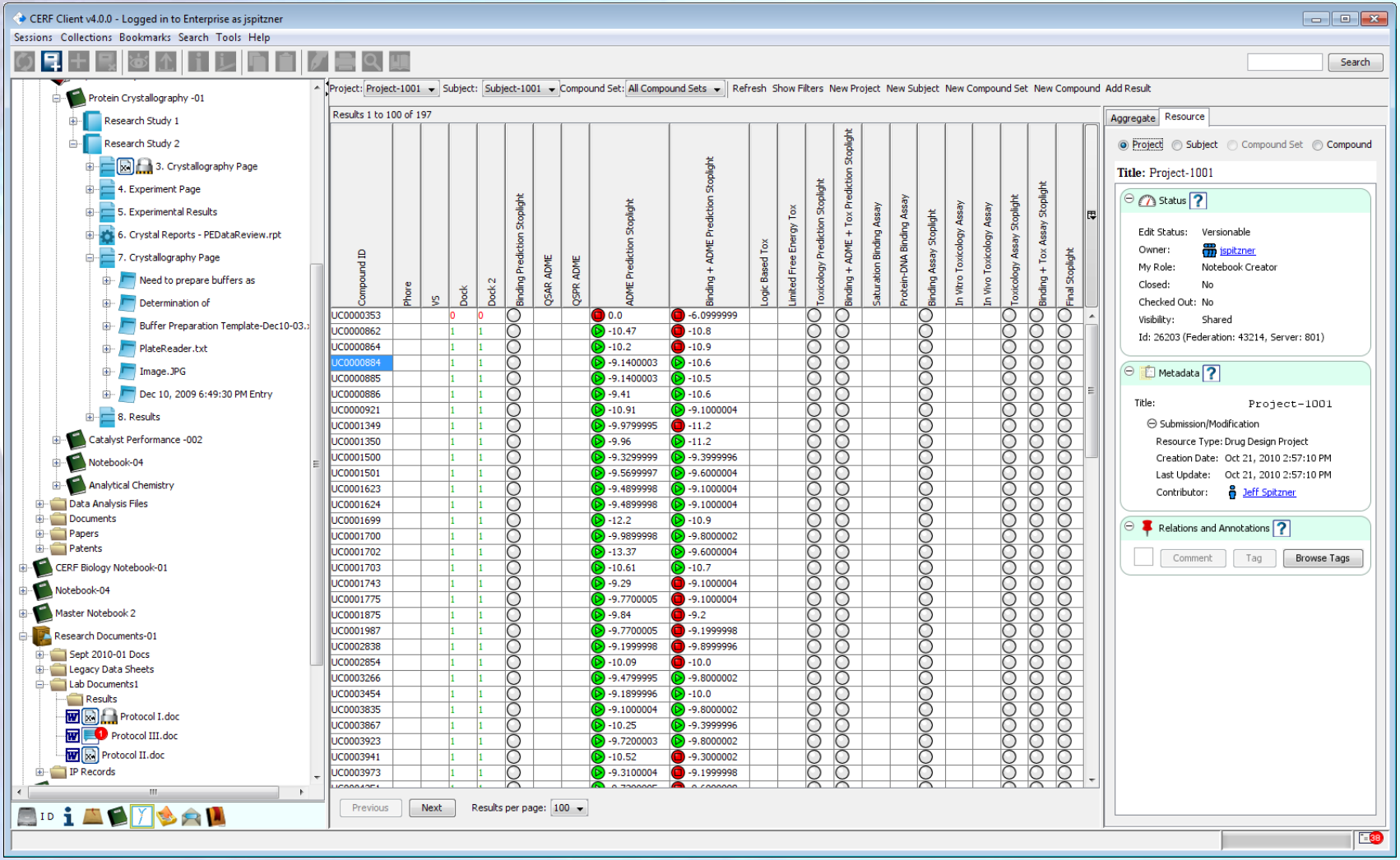
Processing Complex Events Stream





OpenTox

Event Driven Weight of Evidence



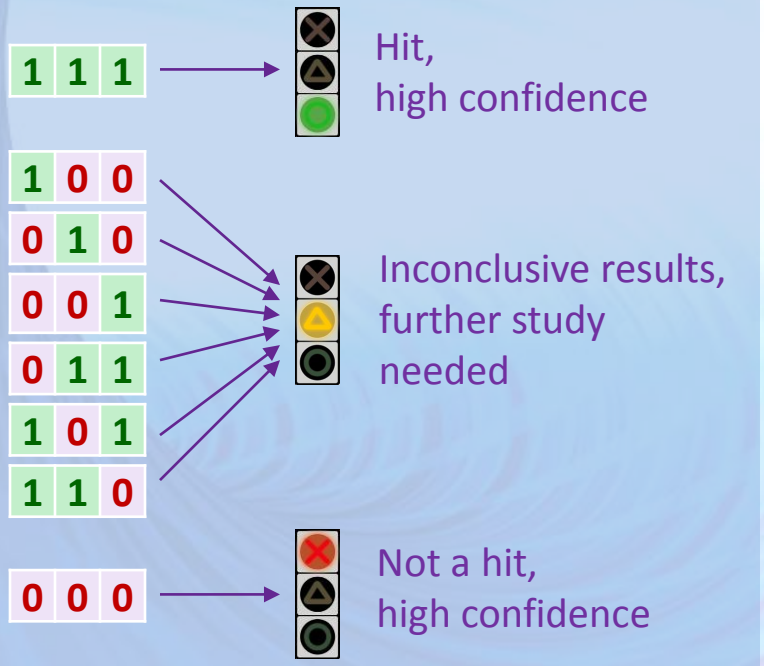
OpenTox

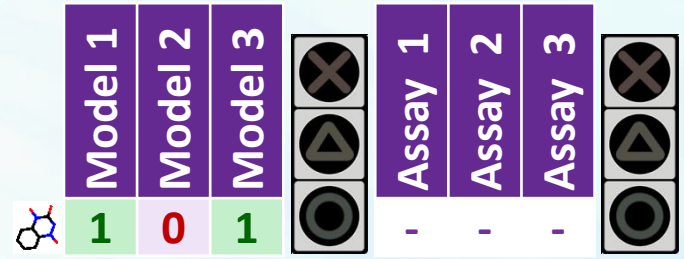
Event Driven Weight of Evidence

collaboration MODERATOR

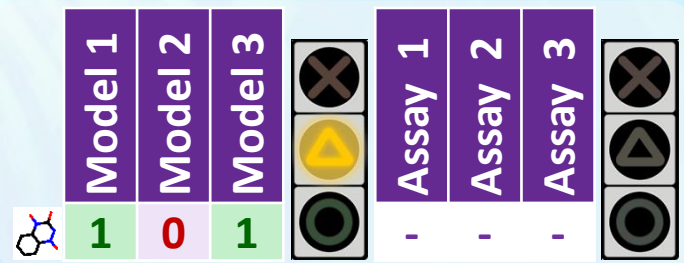
Consensus Rule Editor

Recommendation Rules:

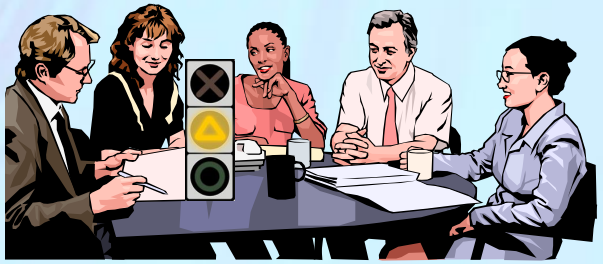




Synergy



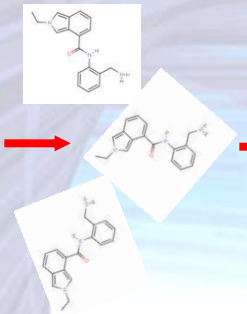
OpenTox



OpenTox

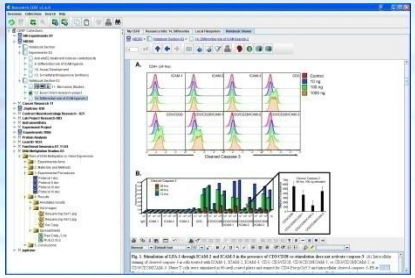
SEVENTH FRAMEWORK PROGRAMME

Synergy Drug Design Collaboration Pilot

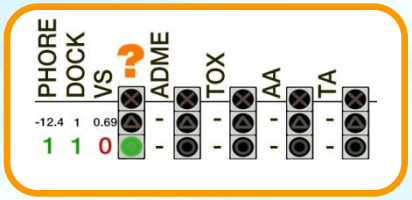


Screened Library

Data

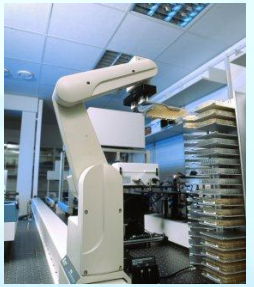


CERF

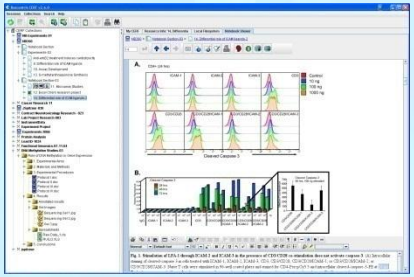


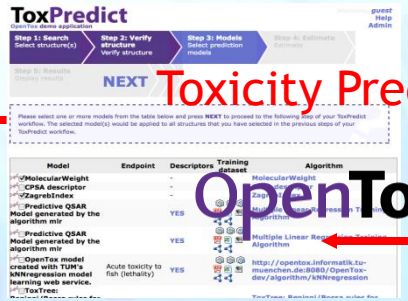
Refine Predictions

SYNERGY



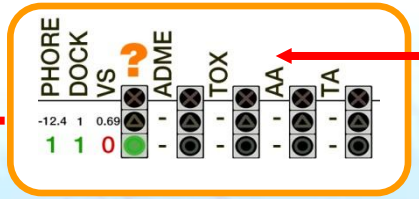
Data



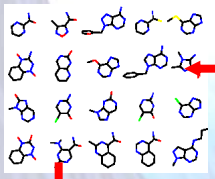


Toxicity Predictions

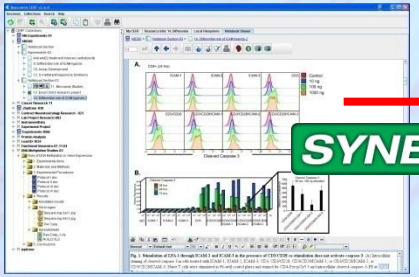
OpenTox



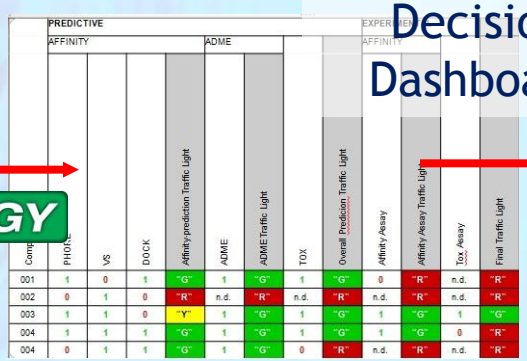
SYNERGY



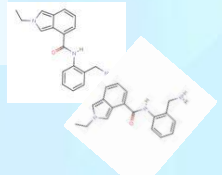
Toxicity Assays



SYNERGY



Decision Dashboard



(Safer) Drug Leads

OpenTox

www.scientistsagainstmalaria.net

SEVENTH FRAMEWORK PROGRAMME

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

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

OpenTox

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.

OpenTox

SEVENTH FRAMEWORK PROGRAMME

Collaborating Partners

Douglas Connect, Switzerland (Coordinator)

In Silico Toxicology, Switzerland

Albert Ludwigs University Freiburg, Germany

Ideaconsult, Bulgaria

National Technical University of Athens, Greece

Istituto Superiore di Sanità, Italy

Fraunhofer Institute for Toxicology & Experimental Medicine, Germany

Technical University of Munich, Germany

David Gallagher, UK

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

Seascape Learning & JNU, India

OpenTox

Our Funding Support...

OpenTox - An Open Source Predictive Toxicology Framework, 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

Contact me:

barry.hardy –(at)- douglasconnect.com

twitter.com/barryhardy

barryhardy.blogs.com

OpenTox