

Ontology Development in support of Predictive Toxicology Use Cases & Services

OpenTox Presentation

16 November 2010

EBI, Hinxton, UK

Barry Hardy (Douglas Connect)



A Predictive Toxicology Pathways Question for you to reflect on during my introduction...

How do we best leverage current knowledge and methods with regards to biological pathway analysis to design improved approaches to predictive toxicology that increase our ability to characterise the potential of chemicals to cause adverse human health effects and including an understanding of mode of action, mechanisms involved in the mode of action and the interaction of biological entities, pathways and networks in the perturbations introduced by the chemicals?

Collaborating Partners

In Silico Toxicology,
Switzerland

Douglas Connect,
Switzerland

Albert Ludwigs University
Freiburg, Germany

Ideaconsult,
Bulgaria

Istituto Superiore
di Sanità, Italy

National Technical
University of Athens,
Greece

Technical University
of Munich, Germany



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

OpenTox Advisory Board

- European Centre for the Validation of Alternative Methods
- Pharmatropé
- Bioclipse
- U.S. Environmental Protection Agency
- U.S. Food & Drug Administration
- Nestlé
- Roche
- AstraZeneca
- LHASA
- Leadscope
- University of North Carolina
- EC Environment Directorate General
- Organisation for Economic Cooperation & Development
- CADASTER
- Bayer Healthcare

Our Funding Support...

For more information on OpenTox,
please visit

www.opentox.org

To join the community site/groups:

www.opentox.org/join_form

Contact me:

barry.hardy -(at)- douglasconnect.com



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).

Journal of Cheminformatics Publication

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

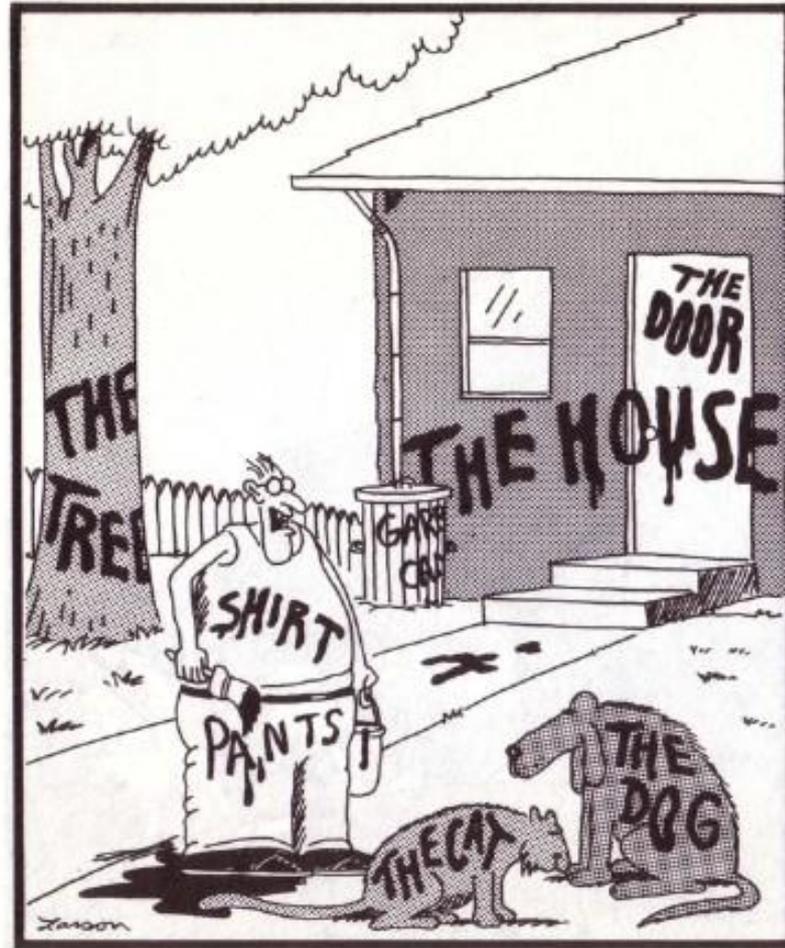
Barry Hardy, Nicki Douglas, Christoph Helma, Micha Rautenberg, Nina Jeliaskova, Vedrin Jeliaskov, Ivelina Nikolova, Romualdo Benigni, Olga Tcheremenskaia, Stefan Kramer, Tobias Girschick, Fabian Buchwald, Joerg Wicker, Andreas Karwath, Martin Gutlein, Andreas Maunz, Haralambos Sarimveis, Georgia Melagraki, Antreas Afantitis, Pantelis Sopasakis, David Gallagher, Vladimir Poroikov, Dmitry Filimonov, Alexey Zakharov, Alexey Lagunin, Tatyana Glorizova, 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 Ontology specific acknowledgements

- Nina Jeliaskova (Ideaconsult)
- OlgaTcheremenskaia and Romualdo Benigni (ISS, Rome)
- Helvi Grimm and Sylvia Escher (Fraunhofer Institute)

Semantic Reflections



"Now! ... *That* should clear up a few things around here!"

Step 1: Search
Select structure(s)

Step 2: Verify structure
Verify structure

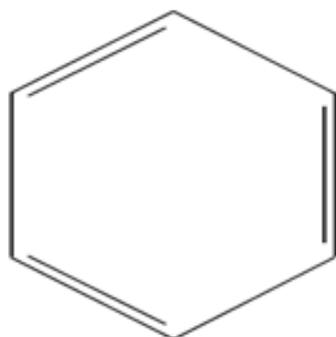
Step 3: Models
Select prediction models

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

Compelling Needs of Users

Integrated Testing

in silico

in vitro

TTC

Read
Across

Category
Formation

REACH Reporting
(QPRF, QMRF)

Applicability
Domain

Validation

Human
Data

Compelling Needs of Users

Multidisciplinary R&D

Good Support of Flexible Applications

Transparency - Not Black Box!

Mechanistic rationale

QSAR & Expert Systems

Workflows

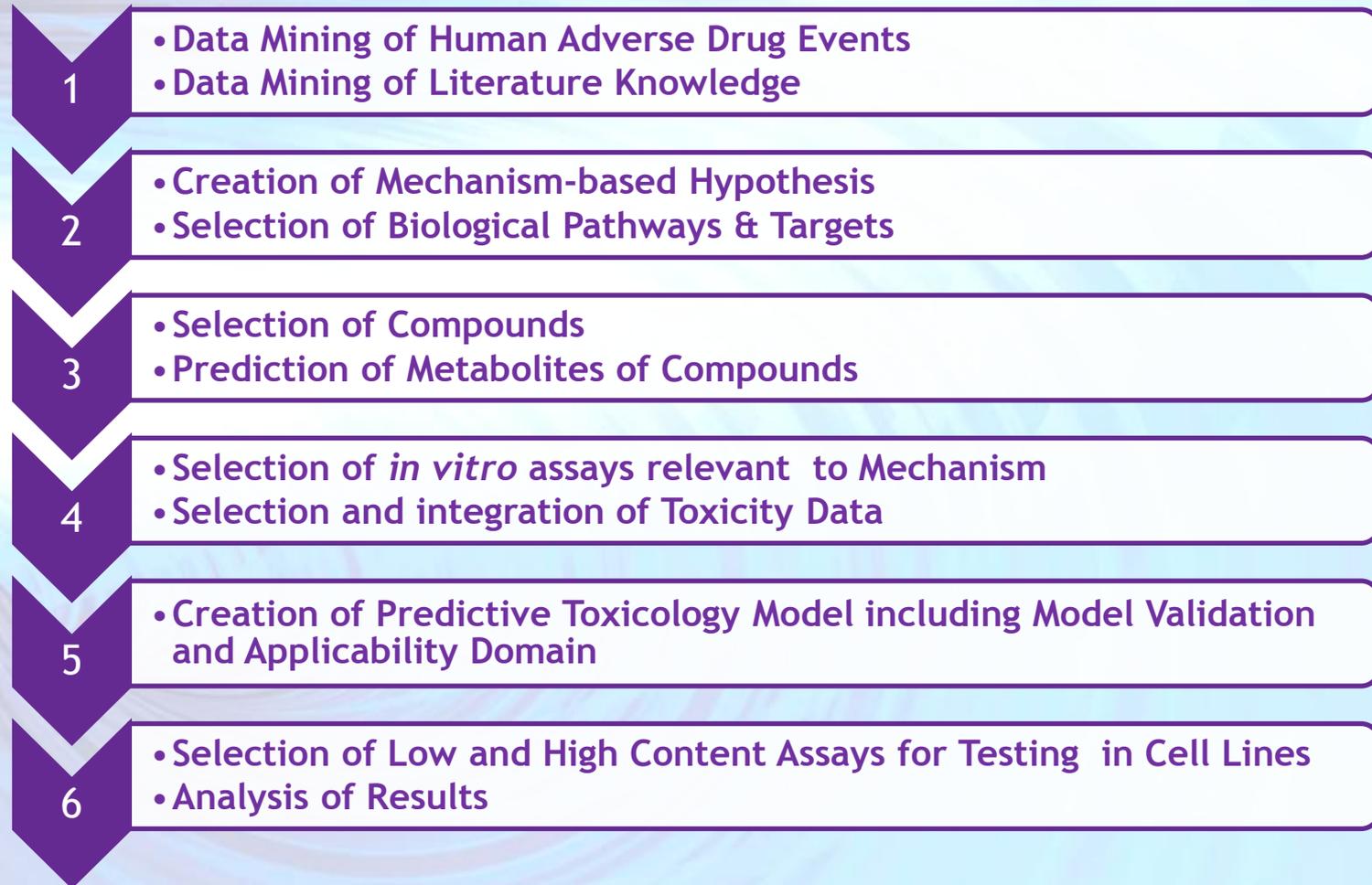
Automated Integration

Applicability Domain

Categories

Systems Biology

OpenTox - Synergy Predictive Toxicology VO Pilot Strategy Development & Case Study



Interacting Components create Solutions



Adaptor Solution in Jeddah, 2008

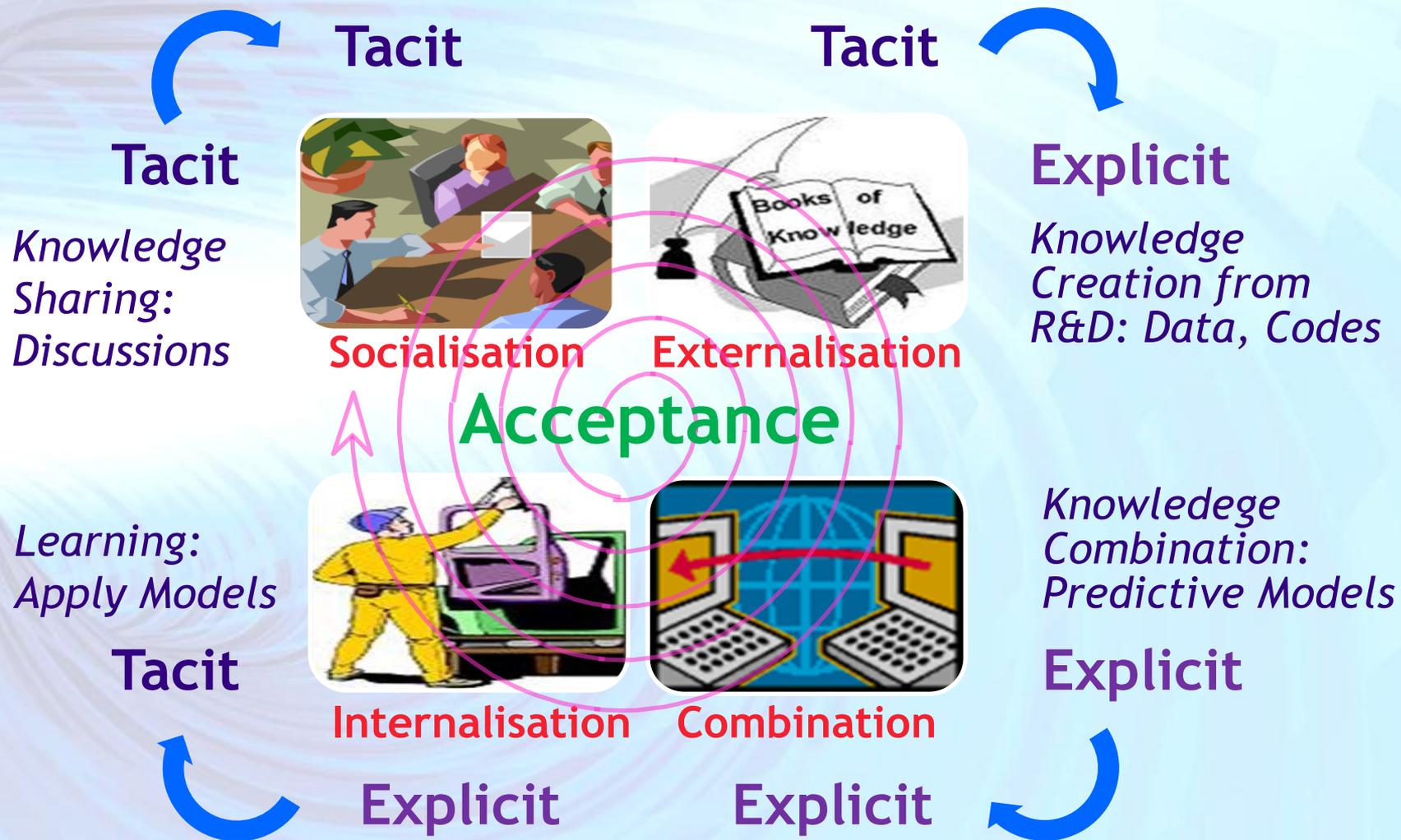
Some Ontological questions for reflection...

- What is knowledge?
- What is a toxicity pathway?
- What is a toxicology ontology?
- Why do we need it?
 - *Knowledge is explicit and tacit*
 - *Accelerating knowledge conversion and flow*
 - *Supporting biological, biochemical, and mechanism based modelling*
 - *Knowledge-based Use Case driven*
 - *Business case*

On defining a knowledge-based approach



SECI Model for Knowledge Management



Complexity Context

Non Repeatable
Adaptative, Patterns,
Filters

Sense
Making for
Emergent
Practice

Leadership
Novel
Practice

Lack of Cause & Effect, Stability-focused
Intervention, Crisis Management



Complex Cause & Effect
Systems Thinking, Analysis

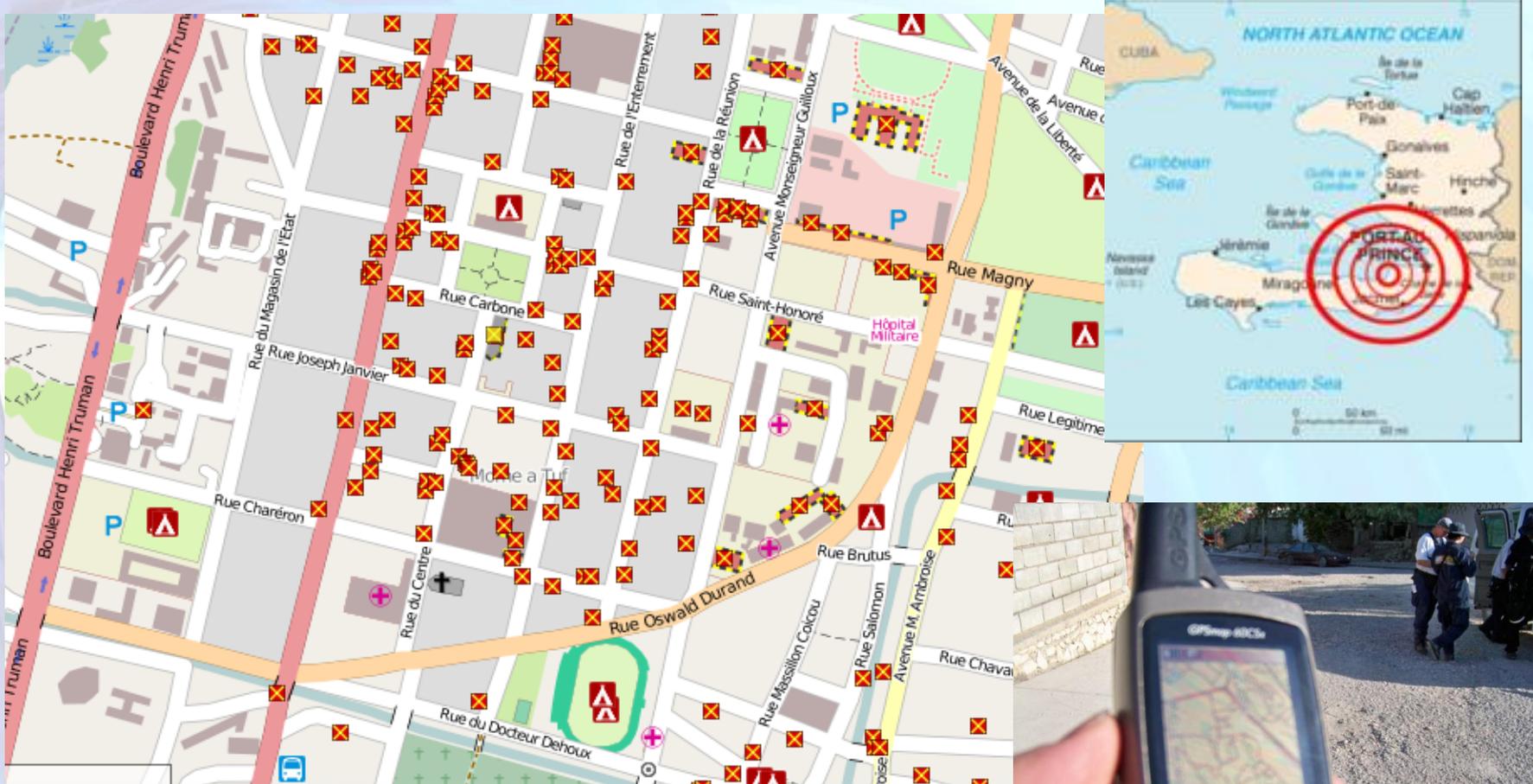
Processes
Good
Practice

Procedures
Best
Practice

Cause & Effect
Repeatable, SOPs



Solution created by Linked Open Data, Web Applications and Crowdsourcing



Haiti Earthquake Crisis Response (2010)

wiki.openstreetmap.org

OpenTox is an Integrating Framework

Framework

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

Diverse Access

- Toxicolog, Biolog, Chem - ists
- Computational Scientists
- Interfaces for new algorithm development & integration

Interoperability

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

	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

OpenTox Components

Compounds: Structures, names, ...

Features: Chemical and biological (toxicological) properties, substructures, ...

Datasets: Relationships between compounds and features

Algorithms: Instructions for solving problems

Models: Algorithms applied to data yield models which can be used for predictions

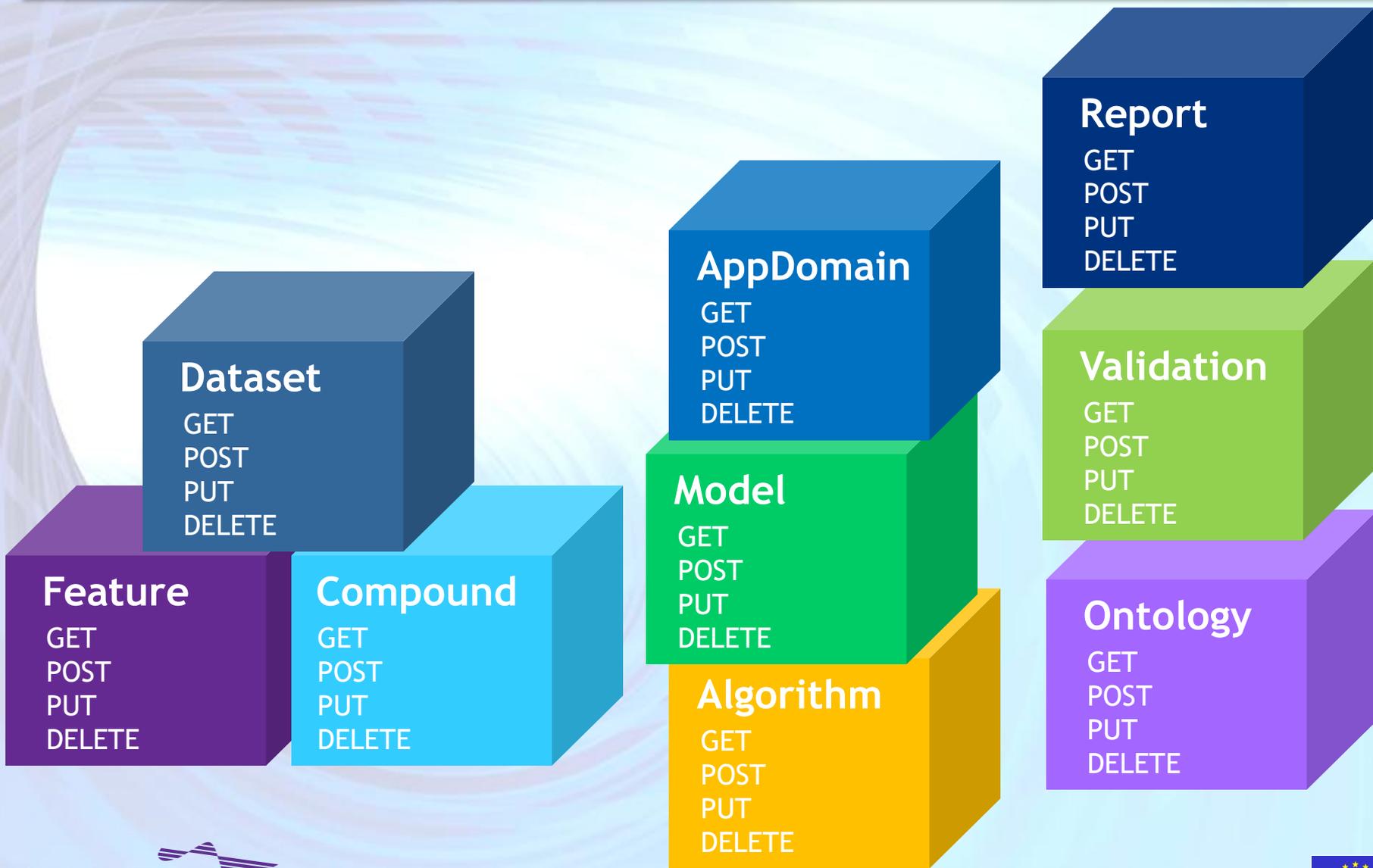
Validation: Methods for estimating the accuracy of model predictions

Reports: Report predictions and models e.g. to regulatory authorities

Tasks: Handle long running calculations

Authentication and Authorisation: Protect confidential data

Overview of Application Programming Interfaces



Representational State Transfer (REST)

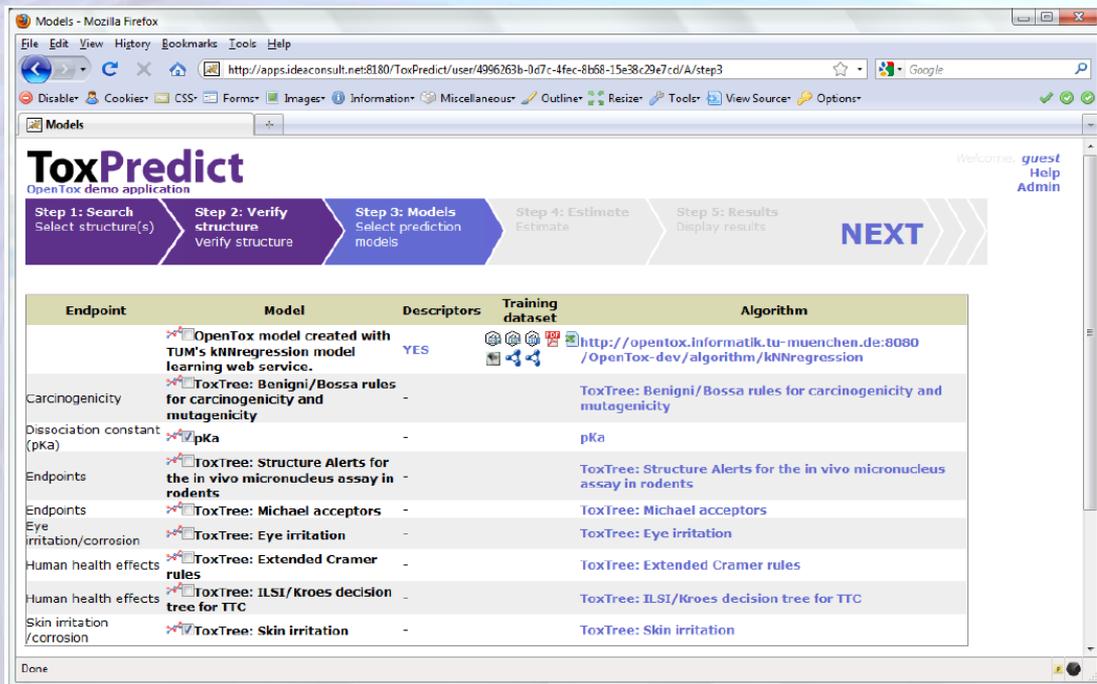
What?

- Architectural style for distributed information systems on the Web
- Simple interfaces, data transfer via **hypertext transfer protocol (HTTP)**, stateless client/server protocol
 - GET, POST, PUT, DELETE
- Each **resource** is **addressed** by its own **web address**

Why?

- **Lightweight** approach to **web services**
- **Simplifies/enables** development of **distributed and local systems**
- Language independent

What you can do with it ...



ToxPredict
OpenTox demo application

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

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

NEXT

Endpoint	Model	Descriptors	Training dataset	Algorithm
Carcinogenicity	OpenTox model created with TUM's kNNregression model learning web service. ToxTree: Benigni/Bossa rules for carcinogenicity and mutagenicity	YES	http://opentox.informatik.tu-muenchen.de:8080/OpenTox-dev/algorithm/kNNregression	ToxTree: Benigni/Bossa rules for carcinogenicity and mutagenicity
Dissociation constant (pKa)	pKa	-		pKa
Endpoints	ToxTree: Structure Alerts for the in vivo micronucleus assay in rodents	-		ToxTree: Structure Alerts for the in vivo micronucleus assay in rodents
Endpoints	ToxTree: Michael acceptors	-		ToxTree: Michael acceptors
Eye irritation/corrosion	ToxTree: Eye irritation	-		ToxTree: Eye irritation
Human health effects	ToxTree: Extended Cramer rules	-		ToxTree: Extended Cramer rules
Human health effects	ToxTree: ILSI/Kroes decision tree for TTC	-		ToxTree: ILSI/Kroes decision tree for TTC
Skin irritation/corrosion	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 interface in a Mozilla Firefox browser window. The URL is <http://apps.ideaconsult.net:8180/ToxPredict/user/4996263b-0d7c-4fec-8b68-15e38c29e7cd/A/step3>. The interface includes a navigation bar with three steps: Step 1: Search (Select structure(s)), Step 2: Verify structure (Verify structure), and Step 3: Models (Select prediction models). Below the navigation bar is a table with columns for Endpoint, Model, and Descriptors.

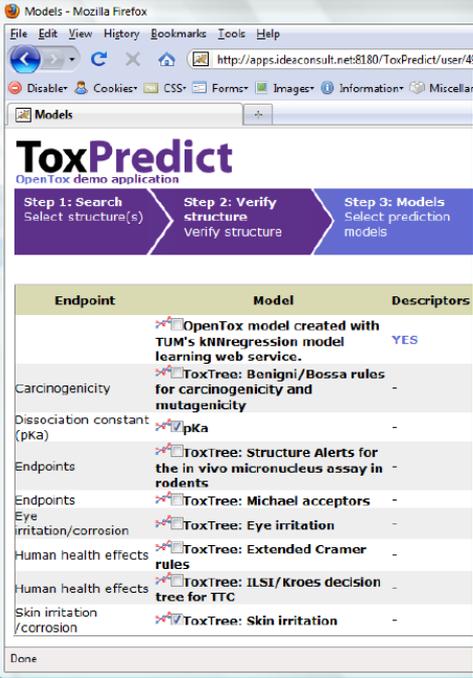
Endpoint	Model	Descriptors
	<input checked="" type="checkbox"/> OpenTox model created with TUM's kNN regression model learning web service.	YES
Carcinogenicity	<input checked="" type="checkbox"/> ToxTree: Benigni/Bossa rules for carcinogenicity and mutagenicity	-
Dissociation constant (pKa)	<input checked="" type="checkbox"/> pKa	-
Endpoints	<input checked="" type="checkbox"/> ToxTree: Structure Alerts for the in vivo micronucleus assay in rodents	-
Endpoints	<input checked="" type="checkbox"/> ToxTree: Michael acceptors	-
Eye irritation/corrosion	<input checked="" type="checkbox"/> ToxTree: Eye irritation	-
Human health effects	<input checked="" type="checkbox"/> ToxTree: Extended Cramer rules	-
Human health effects	<input checked="" type="checkbox"/> ToxTree: ILSI/Kroes decision tree for TTC	-
Skin irritation/corrosion	<input checked="" type="checkbox"/> ToxTree: Skin irritation	-

Below the table, the word "Done" is visible. To the right of the browser window is a map of Europe with five green circular markers placed over various regions.

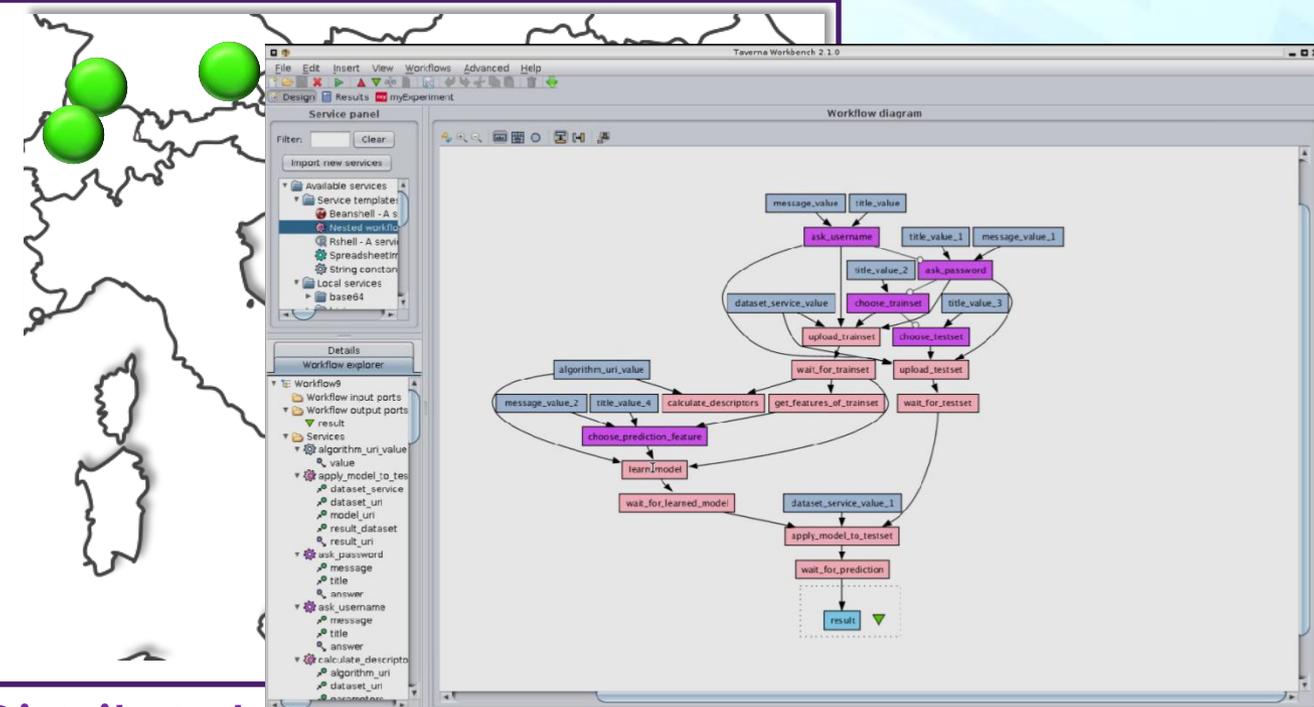
Simple building of applications methods and

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

What you can do with it ...



Endpoint	Model	Descriptors
	OpenTox model created with TUM's kNN regression model learning web service.	YES
Carcinogenicity	ToxTree: Benigni/Bossa rules for carcinogenicity and mutagenicity	-
Dissociation constant (pKa)	pKa	-
Endpoints	ToxTree: Structure Alerts for the in vivo micronucleus assay in rodents	-
Endpoints	ToxTree: Michael acceptors	-
Eye irritation/corrosion	ToxTree: Eye irritation	-
Human health effects	ToxTree: Extended Cramer rules	-
Human health effects	ToxTree: ILSI/Kroes decision tree for TTC	-
Skin irritation/corrosion	ToxTree: Skin irritation	-



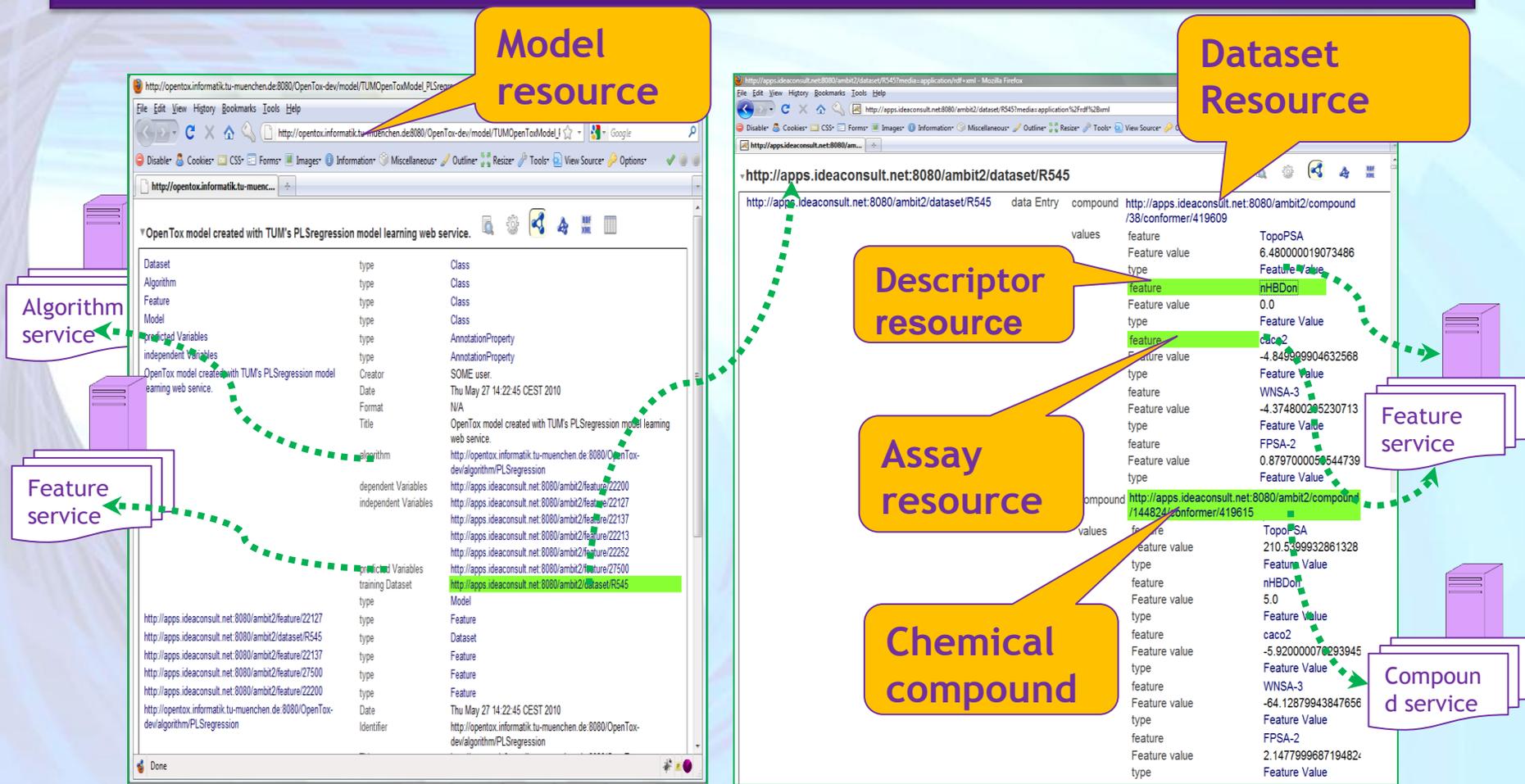
The workflow diagram illustrates a process starting with user input (ask_username, ask_password) leading to dataset selection (choose_trainset, choose_testset), followed by training (upload_trainset, upload_testset, wait_for_trainset, learn_model) and prediction (apply_model_to_testset, wait_for_prediction) steps.

Simple building of applications methods and

Distributed of wide range of methods

Integration into workflow systems for computational biology

Linked resources: Compound, Algorithm, Model, Dataset, Features



Linked resources: Compound, Algorithm, Model, Dataset, Features

Dataset Resource

Descriptor resource

Assay resource

Chemical compound

Browser screenshot: <http://apps.ideaconsult.net:8080/ambit2/dataset/R545>

data Entry	compound	http://apps.ideaconsult.net:8080/ambit2/compound/38/conformer/419609
feature	TopoPSA	
Feature value	6.48000019073486	
type	Feature Value	
feature	nHBDon	
Feature value	0.0	
type	Feature Value	
feature	caco2	
Feature value	8.849999904632568	
type	Feature Value	
feature	WNSA-3	
Feature value	-374800205230713	
type	Feature Value	
feature	FP5A-2	
Feature value	0.8797000050544739	
type	Feature Value	
compound	http://apps.ideaconsult.net:8080/ambit2/compound/144824/conformer/419615	
feature	TopoPSA	
Feature value	210.5399938861328	
type	Feature Value	
feature	nHBDon	
Feature value	5.0	
type	Feature Value	
feature	caco2	
Feature value	-5.920000076293945	
type	Feature Value	
feature	WNSA-3	
Feature value	-64.12879943847856	
type	Feature Value	
feature	FP5A-2	
Feature value	2.147799968719482	
type	Feature Value	

Browser screenshot: <http://apps.ideaconsult.net:8080/ambit2/feature/22213>

Name of the algorithm	type	Class
	type	Class
Numeric Feature	type	Class
	subClassOf	
Source	type	ObjectProperty
Units	type	DatatypeProperty
nHBDon	sameAs	http://www.blueobelisk.org/ontologies/cheminformatics-algorithms/#nHBDonors
Title		nHBDon
Source		http://apps.ideaconsult.net:8080/ambit2/algorithm/org_openscience.cdk.qsar.descriptors.molecular.HBondDonorCountDescr
Units		
type		Numeric Feature

Blue Obelisk algorithms ontology

Regression
Classification
Quantum
Chemistry
Descriptors, etc.

OpenTox algorithm types ontology

Toxicology related ontologies

Browser screenshot: <http://apps.ideaconsult.net:8080/ambit2/feature/22200>

Numeric Feature	type	Class
	type	Class
Source	type	ObjectProperty
Units	type	DatatypeProperty
caco2	Title	caco2
	Source	o049084m_caco2-training_set.sdf
Units	type	
type	sameAs	Numeric Feature
		Gastrointestinal absorption

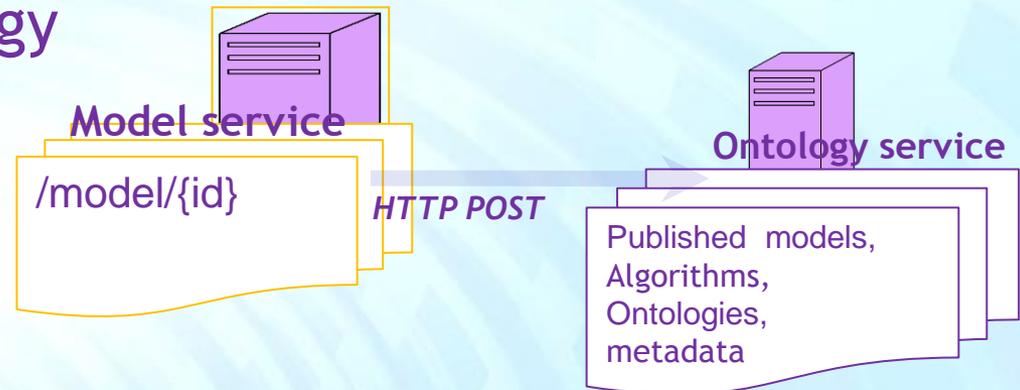
Make the model available

Register at OpenTox ontology service

- RDF triple storage
- Accepts HTTP POST
- SPARQL endpoint

Curl -X POST -d
"uri=http://apps.ideaconsult.net:8080/ambit2/model/57"
<http://apps.ideaconsult.net:8080/ontology>

Becomes visible for applications



The screenshot shows the ToxPredict web application interface. It features a progress bar with five steps: 1. Select structure(s), 2. Verify model(s), 3. Select model(s), 4. Run (ToxService), and 5. Display results. A 'NEXT' button is visible at the end of the progress bar. Below the progress bar is a table with columns: Model, Endpoint, Algorithm, and Validation. The table contains several rows of data, including models like 'ToxTree: Verhaar scheme for predicting toxicity mode of action' and 'ToxTree: Benigni/Bossa rules for carcinogenicity and mutagenicity'.

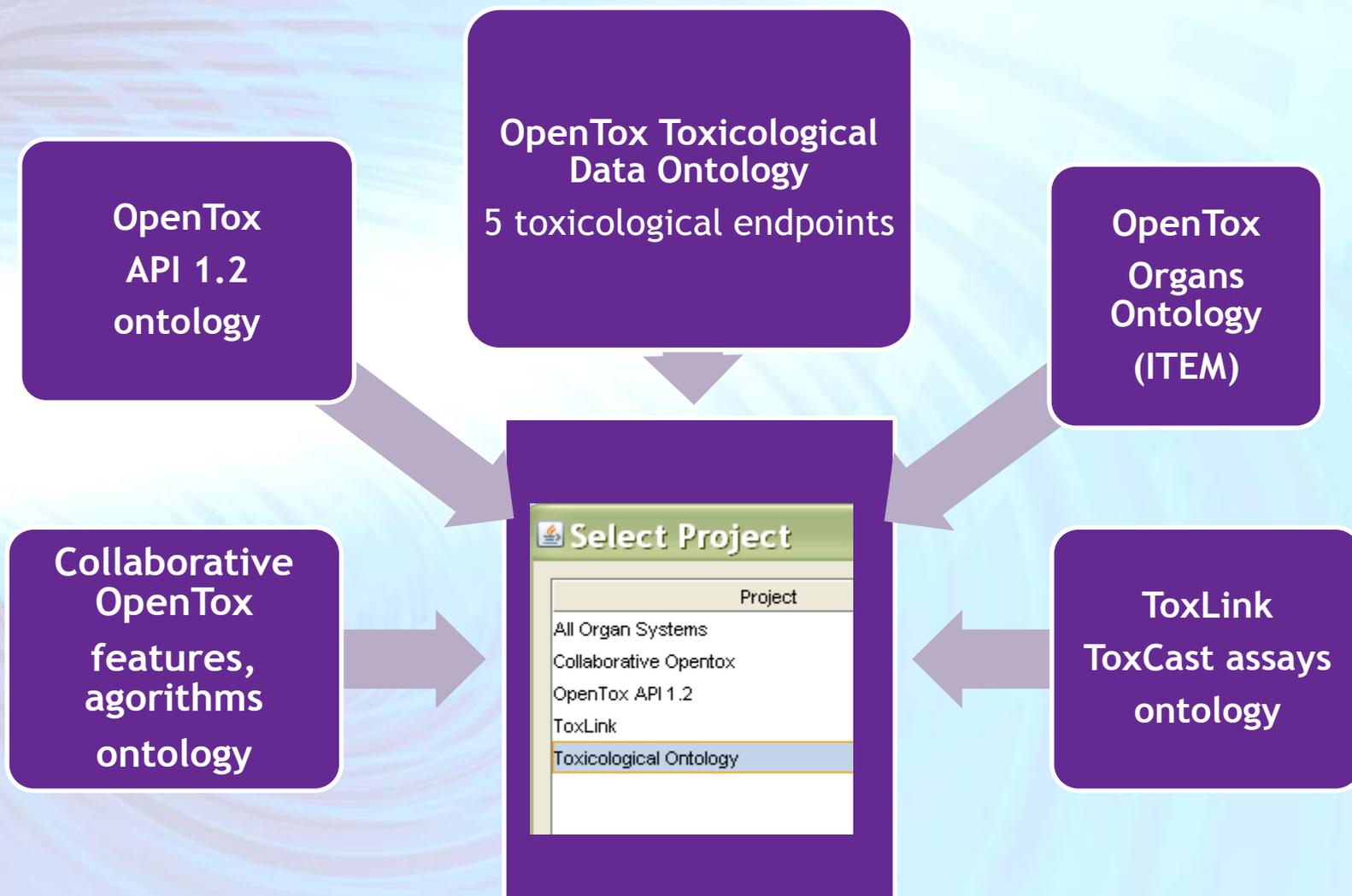
Model	Endpoint	Algorithm	Validation
CholesterolWeight		MolecularWeight	
ToxTree: Verhaar scheme for predicting toxicity mode of action	Acute toxicity to fish (lethality)	ToxTree: Verhaar scheme for predicting toxicity mode of action	
ToxTree: Benigni/Bossa rules for carcinogenicity and mutagenicity	Carcinogenicity	ToxTree: Benigni/Bossa rules for carcinogenicity and mutagenicity	
LogKa	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/commission	ToxTree: Eye Irritation	
Caco-2 Cell Permeability	Gastrointestinal absorption	Regression: Linear regression	Model validation report
OpenTox model created with TUM's PL Segregation model learning web service.	Gastrointestinal absorption	http://opentox.infomark.it/monoschenko/0000/OpenTox_dev/algorithms/PL_Segregation	
OpenTox model created with TUM's MN-Regression model learning web service.	Gastrointestinal absorption	http://opentox.infomark.it/monoschenko/0000/OpenTox_dev/algorithms/MN-Regression	
Lipinski Rule of Five	Human health effects	Lipinski Rule of Five	
ToxTree: Cramer rules	Human health effects	ToxTree: Cramer rules	
ToxLogP	Octanol-water partition	ToxLogP	

Need for communications in the community overcoming different languages and vocabularies



Explaining the rules of different games on a conservation project trip in the Caprivi, Namibia

Collaborative Ontology Development: Collaborative Protege Server



ToxML: conversion in OWL ontology

- Initial work to check out how costly would be to convert ToxML to an OWL ontology
- Why:
 - -to integrate the Leadscope databases in the OpenTox service
 - -to combine with data coming from different databases (e.g. for complex queries)

ToxML
xsd
schema



Protégé OWL

File Edit Project OWL Reasoning

Metadata(Toxicity) OWLClass

SUBCLASS EXPLORER

For Project: ● CarcinogenicityStudies

Asserted Hierarchy

- owl:Thing
 - ts:CData
 - ▶ ● ts:Compound
 - ts:ControlType
 - ts:InexactValue
 - ts:Quantity
 - ts:RegulatoryTestType
 - ts:ResultFindings
 - ts:Sex
 - ts:TestCall
 - ts:TypedValue

Some parts of the
taxonomy may need to
be reorganized

Procedures:

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

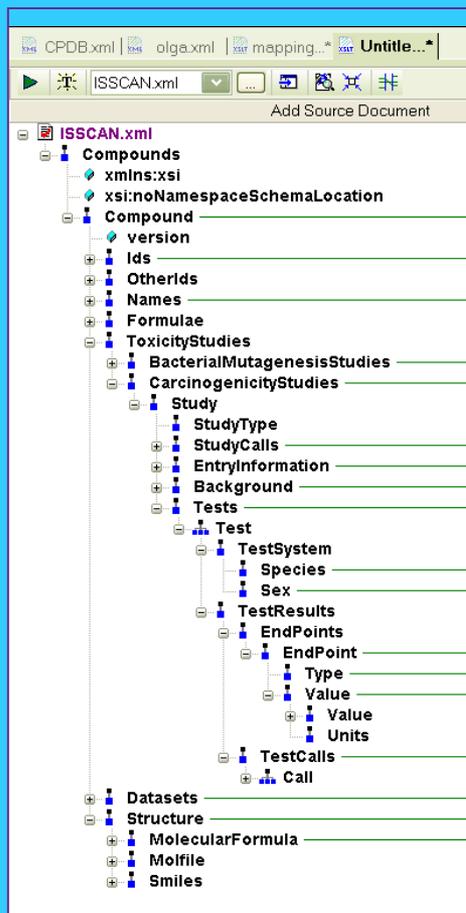
each element which has a type in the summary file is considered as a property

the parent relationships are kept

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

Toxicological Endpoint Ontology Development

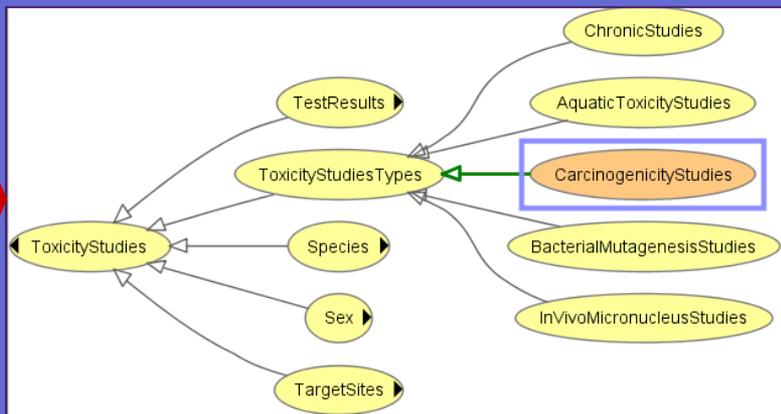
ToxML schema



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

OpenTox
Toxicological
Endpoint
Ontology

Ontology Development

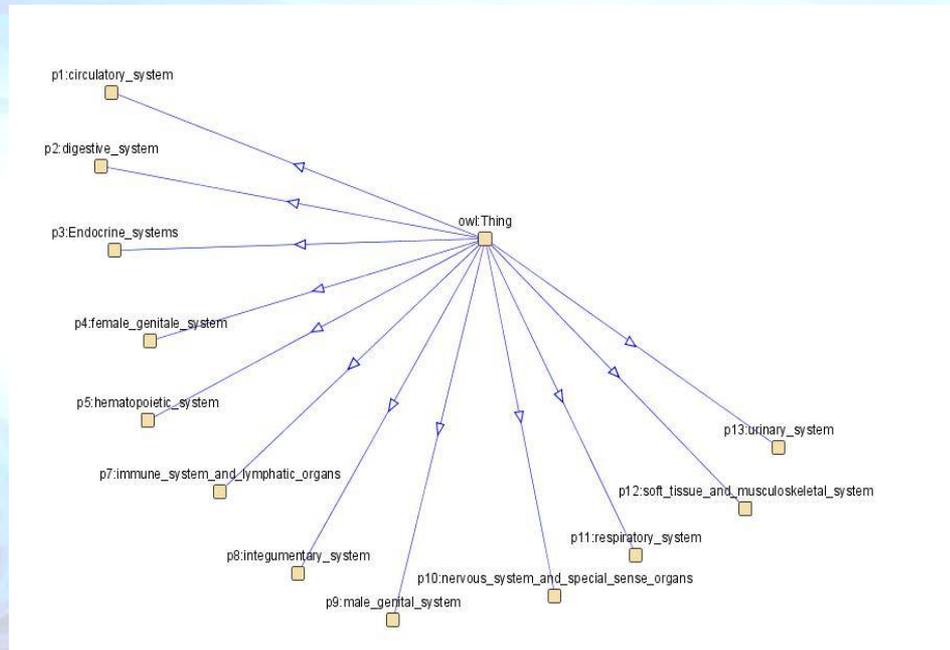


Re-use of terms defined in
neighbouring ontologies (e.g. OBO)

Collaborative
Protégé
Environment

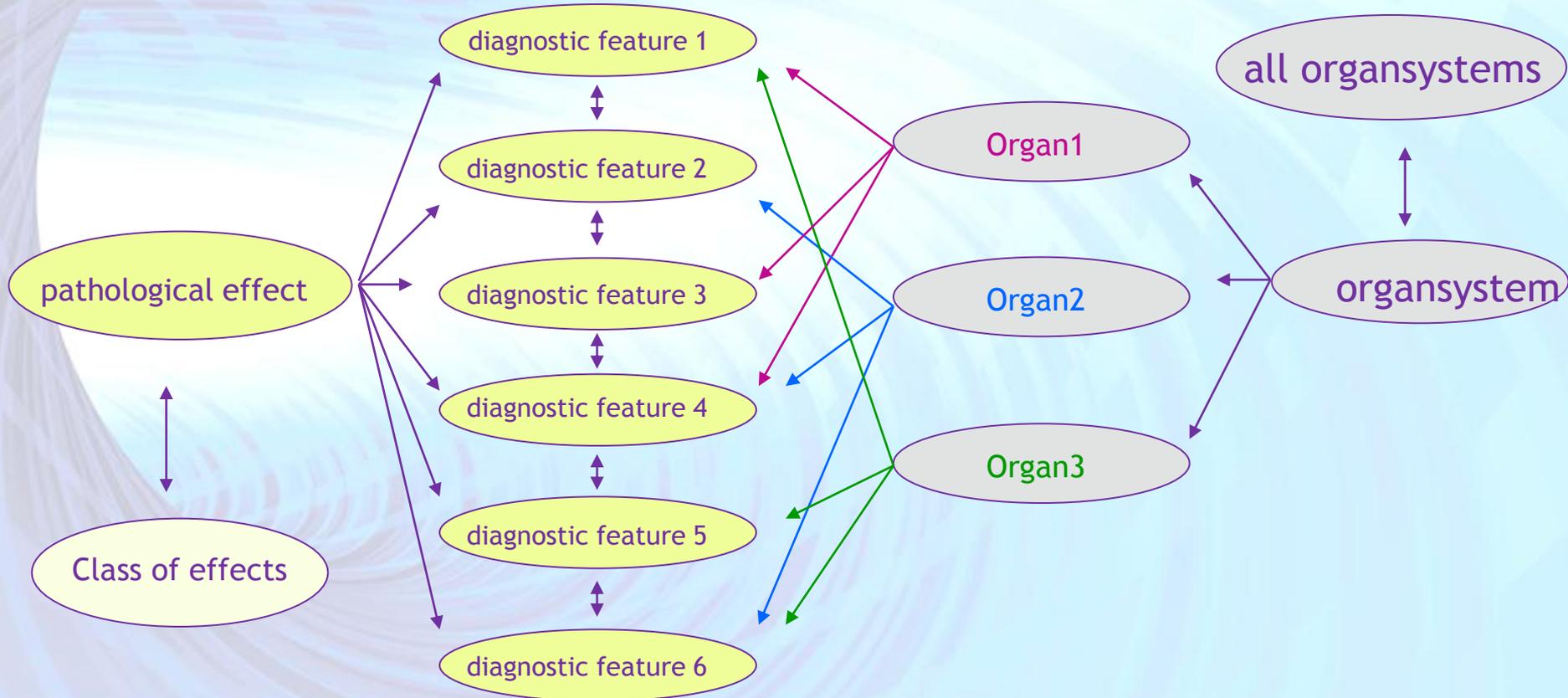
OpenTox Organ Ontology Development

- organ ontology consisting of 12 very detailed organsystems



- effect ontology, which is linked to the organ ontology
- comprehensive review by FhG pathologists, who have been involved in the INHAND process

OpenTox Organ Ontology



ToxLink: ToxCast Ontology

ToxLink Protégé 3.4.3 (rmi://opentox/ToxLink)

File Edit Project OWL Reasoning Code Tools Window Collaboration Help

Metadata(EPANCCTAssayOntology1.owl) OWLClasses Properties Individuals

SUBCLASS EXPLORER

For Project: ●

Asserted Hierarchy

- owl:Thing
 - Assay
 - AssayComponent
 - AssayParameter
 - AssayCode
 - AssayReadout
 - AssayTarget
 - AssayTechnology
 - BioEffect
 - BioInteraction
 - BioProcess
 - Disease
 - Phenotype
 - BioTissue
 - Cell
 - Organ
 - Organism
 - Tissue
 - GeneBag
 - GOCategory
 - Pathway

CLASS EDITOR for Assay (instance of owl:Class)

For Class: <http://www.semanticweb.org/ontologies/2010/4/EPANCCTAssayOntology1.owl#Assay>

Property	
rdfs:comment	

owl:Thing

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

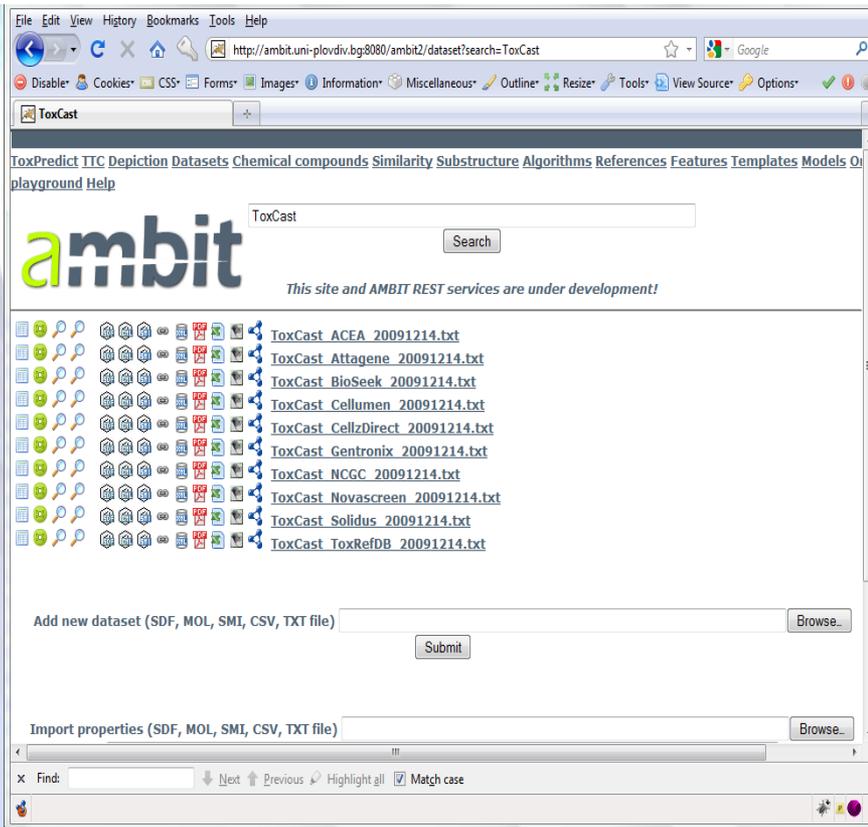


www.opentox.org/opentoxipedia



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
```

Example: ToxCast

```

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#>

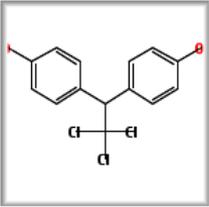
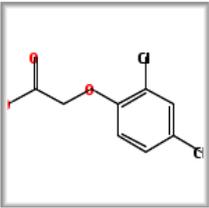
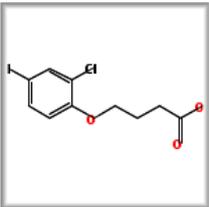
select ?Feature ?title ?id ?assay ?geneid ?gene
where {
  ?Feature rdf:type ot:Feature.
  {?Feature dc:title ?title}.
  {?Feature owl:sameAs ?assay}.
  {?assay toxcast:gene ?geneid}.
  {?geneid toxcast:hasProperty ?genename}.
  {?genename rdf:type toxcast:GENE_NAME}.
}
    
```



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

Chemical compounds

Search results Dataset = 961 hits: 100

#	Compound	ToxCast At	Benigni /	Benigni /
		ATG RORE CIS	Structural Alert for genotoxic carcinogenicity	Structural Alert for nongenotoxic carcinogenicity
1		1000000.0	NO	NO
2		1000000.0	NO	NO
3		1000000.0	NO	NO
4		1000000.0	NO	NO

SPARQL

```

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

Submit Query

Results [found]

Feature

- <http://ambit.uni-plovdiv.bg:8082/ontology/feature/335126>
- <http://ambit.uni-plovdiv.bg:8082/ontology/feature/335127>
- <http://ambit.uni-plovdiv.bg:8082/ontology/feature/335132>
- <http://ambit.uni-plovdiv.bg:8082/ontology/feature/335133>
- <http://ambit.uni-plovdiv.bg:8082/ontology/feature/335134>
- <http://ambit.uni-plovdiv.bg:8082/ontology/feature/335135>
- <http://ambit.uni-plovdiv.bg:8082/ontology/feature/335136>

Find: Done

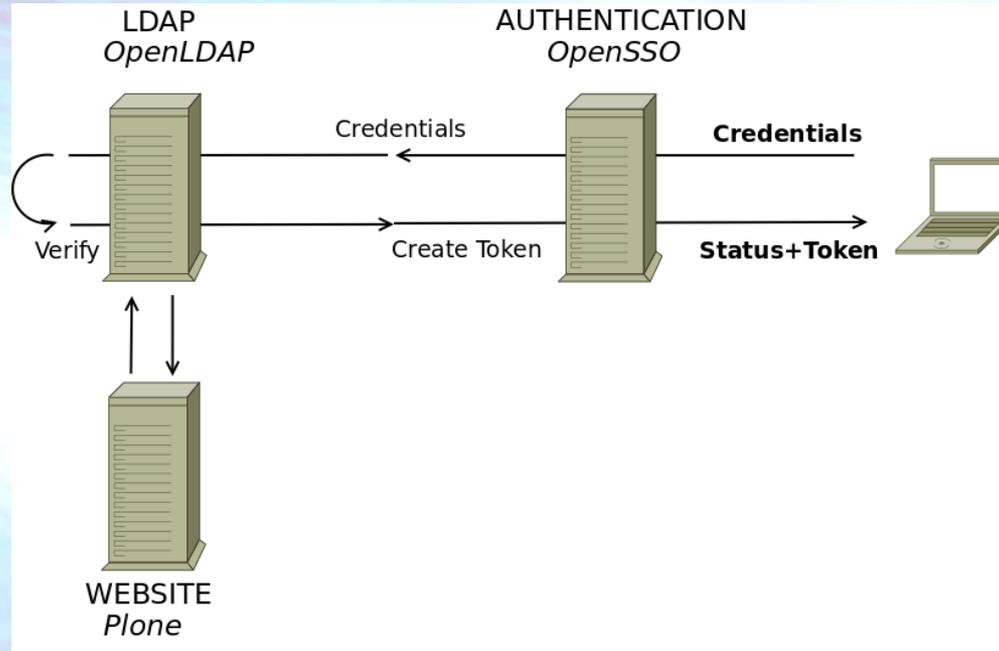
?feat

126

Controlling Access to Confidential Information

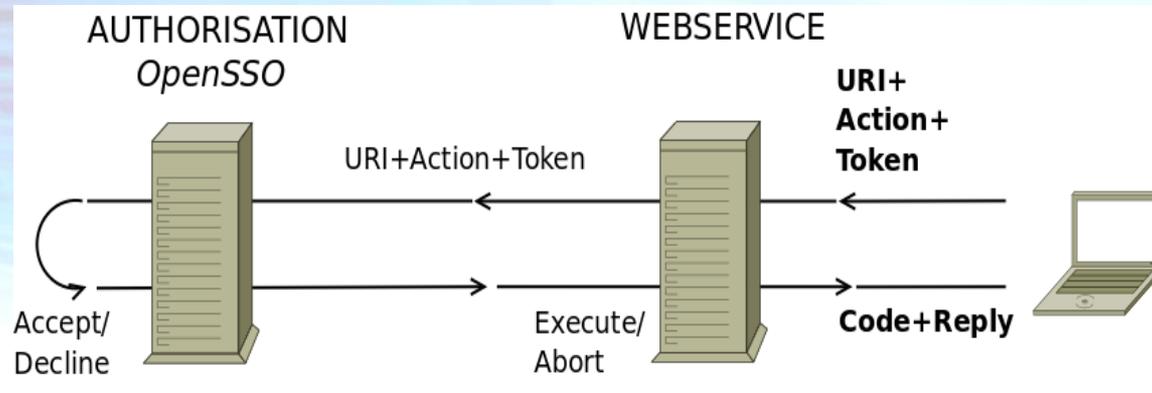
- OpenTox makes resources available through **URIs**
- OpenTox provides facilities to protect confidential information located at **URIs**. Two tasks are involved here:
 - **Authentication**: Confirming the identity of the user requesting access
 - **Authorisation**: Granting the confirmed identity access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



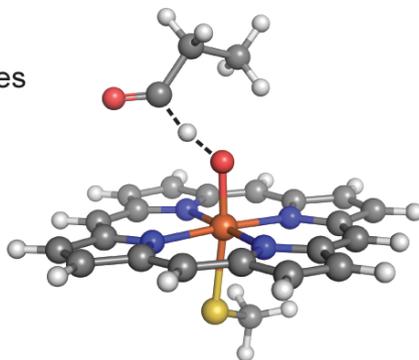
- Tokens encode user identity
- Tokens are valid for a certain time period only (customizable)
- The triplet URI+Action+Token makes up the call to be authorised
- All messages are encrypted (SSL)
- Resource Owners create and modify policies defining access rules

SMARTCyp Service for Predicting Metabolites

Atom Reactivity Library

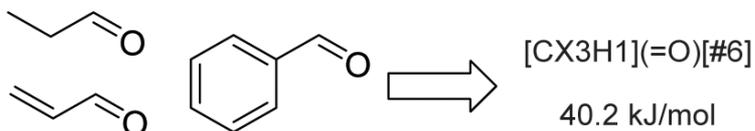
A. Calculate Quantum Chemical Reference Energies

Calculate transition state energies using density functional theory



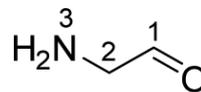
B. Define SMARTS Rules

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

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



3. Compute Score and Rank Atoms

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

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

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

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

Atom 1 - Rank 2

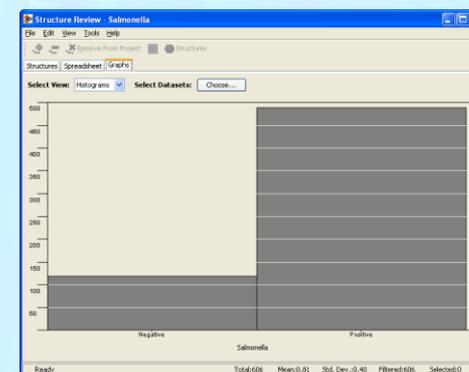
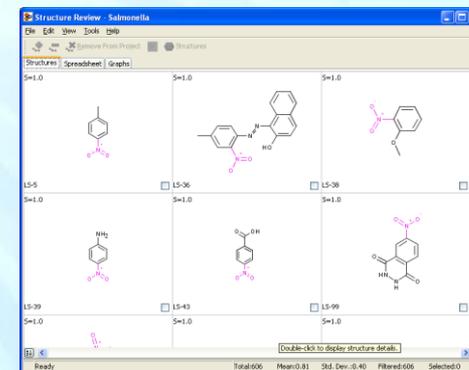
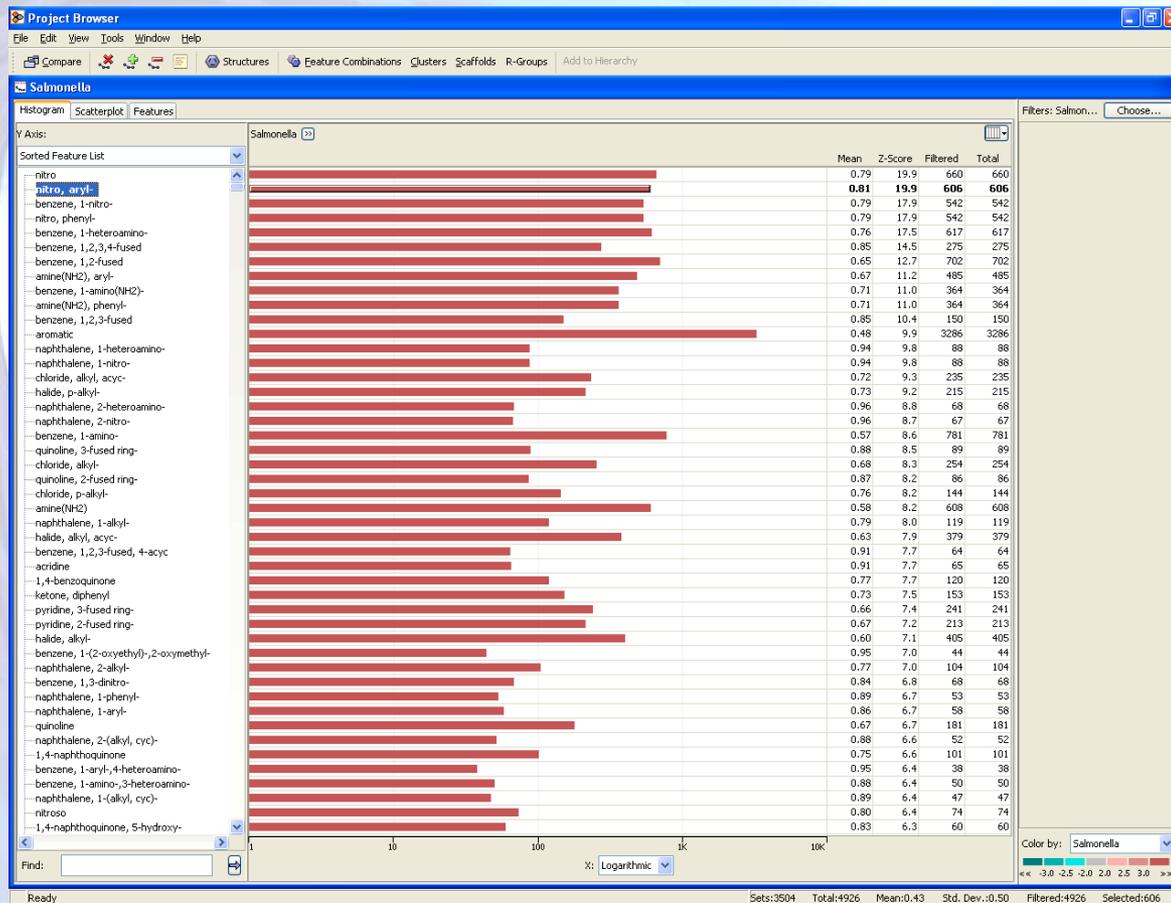
Atom 2 - Rank 1

Atom 3 - Rank 3

SMARTCyp - developed by Patrik Rydberg, University of Copenhagen

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

OpenTox - Leadscope



OpenTox - Bioclipse



UPPSALA
UNIVERSITET

Problem

Building
Blocks

Conclusion

Bioclipse

The screenshot displays the Bioclipse application window. On the left is the Bioclipse Navigator showing a project tree with folders like 'CDKWS2009', 'ChemGate', 'Basis', 'Media', 'NMRshiftool', 'OWL', 'PDF', 'SampleData', 'SampleData', 'Solubility', 'STATLite', and 'Test'. The 'Test' folder is expanded to show files like 'bug824.js', 'foo.js', 'Fragments2_3d.sdf', 'gist.22185.js', and several 'unnamed.cml' files. The main workspace shows a table with two rows of molecular data:

	2D-structure	MOLREGNO
14577		17159
14578		17160

On the right, the Properties window shows the 'General' tab with the following data:

Property	value
Has 2D Conds	yes
Has 3D Conds	no
Molecular Format	N/A
Molecular Formula	C18H26H26N06P
Molecular Mass	358.1781
Molecular Properties	
MOLREGNO	17159

O. Spjuth et al., *BMC Bioinformatics* 2007, 8:59



UPPSALA
UNIVERSITET

Problem

Building
Blocks

Conclusion

The Chemistry Development Kit

A Family of Projects

- CDK-Taverna (chemoinformatics workflows)
- JChemPaint (semantic 2D editor)
- ChemoJava (GPL-ed extension)

Goals

- library of cheminformatics algorithms
- educational

Usage

- CDK: 100+ times cited in scientific literature
- Bioclipse, KNIME, Jumbo (CML), AMBIT, ...

C. Steinbeck et al., J.Chem.Inf.Comput.Sci, 2003

C. Steinbeck et al., Curr.Pharm.Design, 2006

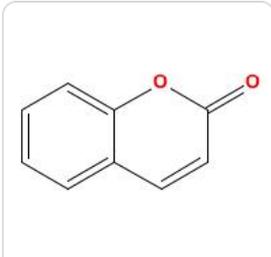
Bioclipse Visualisation Workbench

The screenshot displays the Bioclipse Decision Support interface. The main window shows the assessment results for the compound coumarin. The interface includes a toolbar at the top, a report editor window, and a decision support panel on the right.

Bioclipse Decision Support
Chemical Liability Assessment

Report date: 2010/08/26

Query structure:



Properties

Mol.	146.143
H donors:	1
H	1
alogP:	1.031

Number of Consensus: 4
Endpoints: 2 negative, 1 inconclusive

Compound: coumarin
SMILES: O=C1OC2=CC=CC=C2C=C1
Formula: C9H6O2
InChI: InChI=1S/C9H6O2/c10-9-6-5-7-3-1-2-4-8(7)/11-9/h1-6H

Endpoint: AHR ✓ **NEGATIVE**

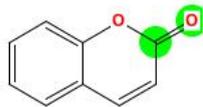
Model: AHR exact matches
Consensus: INCONCLUSIVE

Model: AHR nearest neighbour
Consensus: ✓ NEGATIVE

Model: AHR Signature Alerts
Consensus: ✓ NEGATIVE

Model: AHR Signature Significance
Compound: [O](=[C])
Classification: **NEGATIVE**
Consensus: ✓ NEGATIVE

Details:



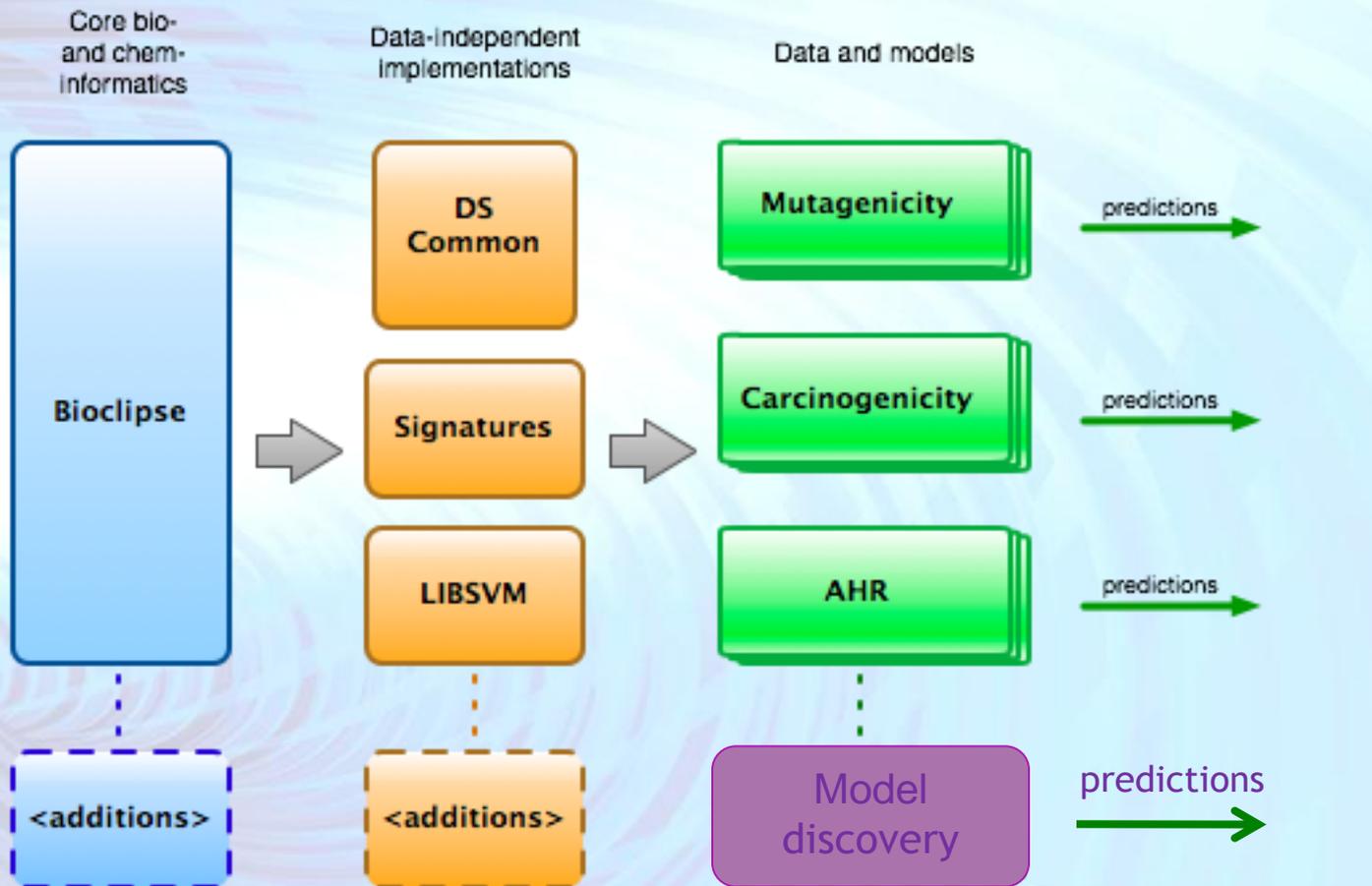
1 of 3

Decision Support Panel:

- AHR
 - AHR Signature Alerts
 - AHR Signature Significance [1 neg]
 - [O](=[C])
 - AHR exact matches
 - AHR nearest neighbour
 - Carcinogenicity
 - CPDB Signature Alerts
 - CPDB Signature Significance [1 pos]
 - Result: 1.002
 - CPDB exact matches [1 pos]
 - Index 199
 - CPDB nearest neighbour
 - Mutagenicity
 - Ames Signature Significance [1 neg]
 - [C]([C][O]=[O])
 - Ames Structural Alerts
 - Ames exact matches [1 pos]
 - 91-64-5
 - Ames nearest neighbour [3 neg]
 - 90-33-5 [tanimoto=0.78]
 - 2107-76-8 [tanimoto=0.76]
 - 26093-31-2 [tanimoto=0.75]

Consensus: **NEGATIVE** ✓

Bioclipse - OpenTox Interoperation



OpenTox - ToxCast

ACToR: Aggregated Computational Toxicology Resource
U.S. ENVIRONMENTAL PROTECTION AGENCY

Data Collection: EPA CCL3

Name: EPA CCL3 [List_001](#)

Description: EPA has drinking water regulations for more than 90 contaminants. The Safe Drinking Water Act (SDWA) includes a process that we must follow to identify and list unregulated contaminants which may require a national drinking water regulation in the future. EPA must periodically publish this list of contaminants (called the Contaminant Candidate List or CCL). In February 2008 we announced the draft CCL 3.

ID: 139

Institutional Source: EPA

Source Type: Chemicals

Number of Substances: 93

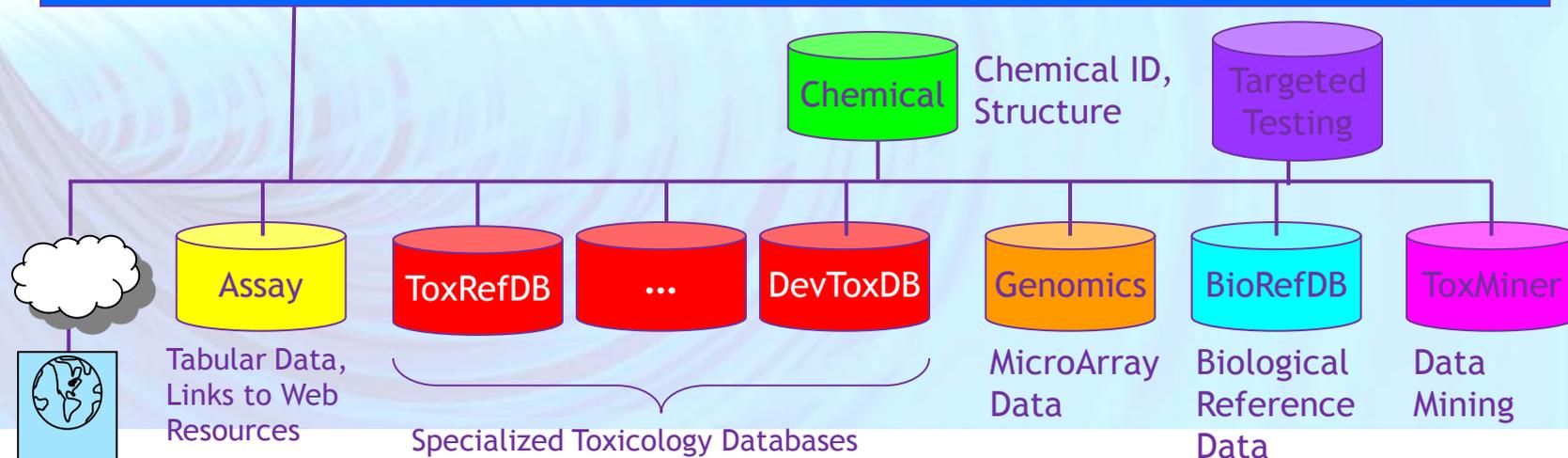
Number of Generic Chemicals: 92

Chemical Table
Page 1 of 2: [Load](#)

Structure	Name	CASRN	Generic Chemical Details	Toxicity Endpoints					
				Hazard	Chronic Toxicity	Developmental Toxicity	Genotoxicity	Reproductive Toxicity	Chemical Toxicity
<chem>ClC(Cl)(Cl)Cl</chem>	1,1,1,2-Tetrachloroethane	630-20-6	Details	Ha	Ca	G	D	R	Cr
<chem>CC(=O)Cl</chem>	1,1-Dichloroethane	75-34-3	Details	Ha	Ca	G	D	R	Cr

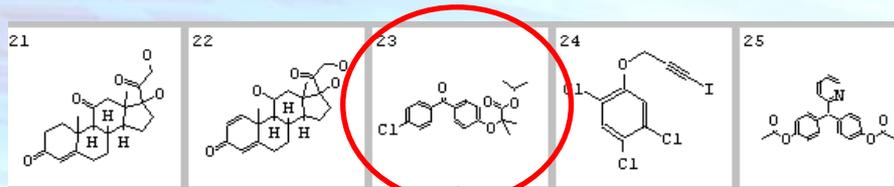
ACToR Web Browser

ACToR API

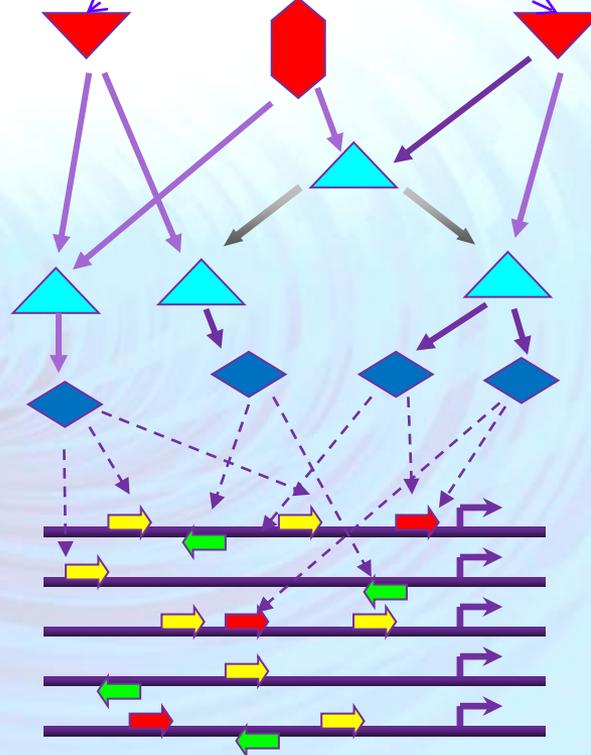


actor.epa.gov

Integrated Workflow of Bioinformatics, Systems Biological and Cheminformatics Tools



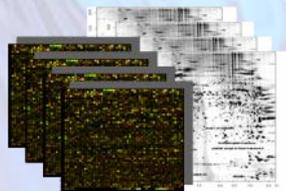
SAR/QSAR



Cheminformatics
Biological activities of the compounds

Systems Biology
Network analysis and simulation, potential drug targets

Bioinformatics
Regulatory modules, transcription factors



Lets talk in:

BioUML

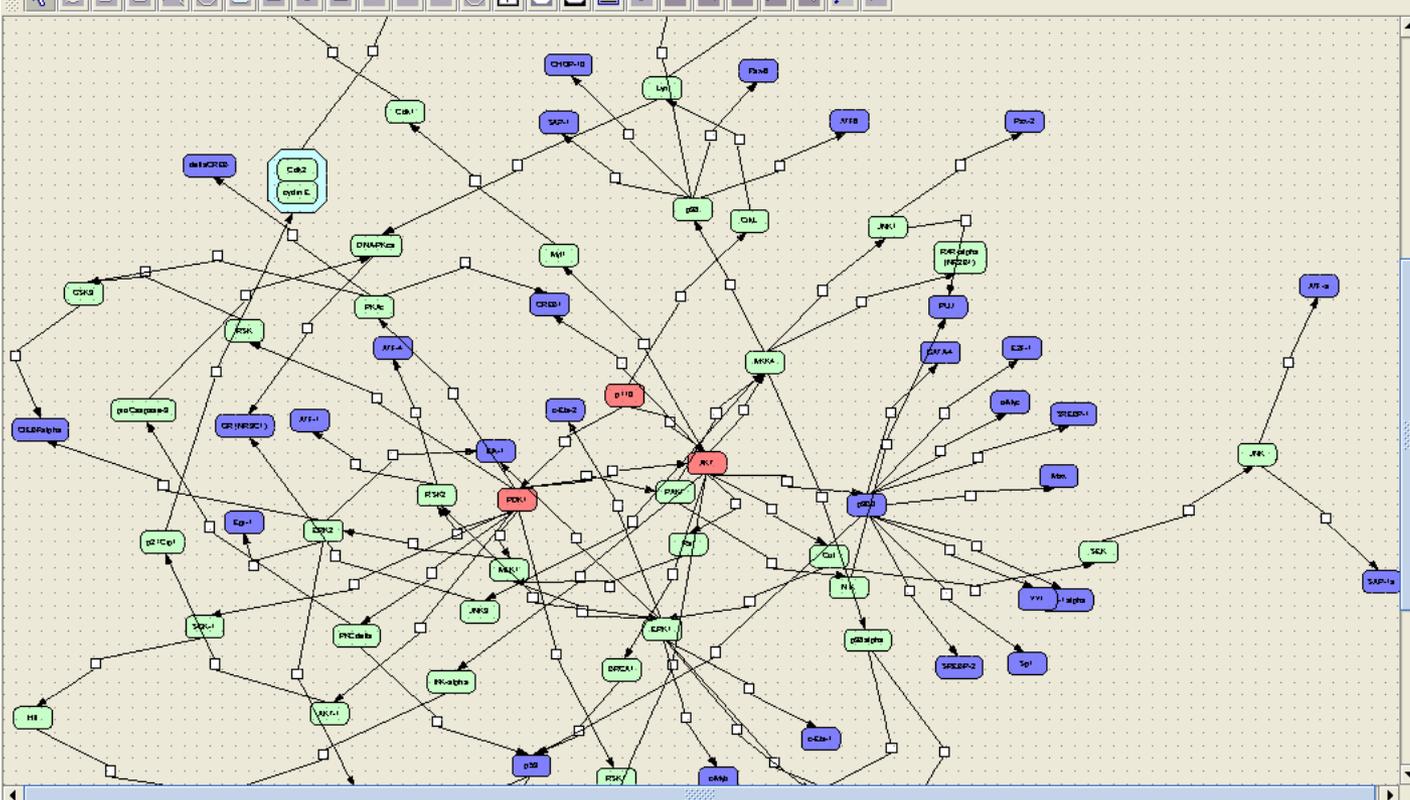
Unified Modeling Language (UML) is a standardized general-purpose modeling language



Databases Data Analyses

- data
 - Collaboration
 - Morgun
 - Morgun2
 - RITA
 - Data
 - Diagrams
 - Drug_target
 - Files
 - Scripts
 - TFsites_nonredun
 - TFsites_nonredun
 - TFsites_nonredun
 - Tables
 - 3DT
 - 4DT
 - PASS_molecu
 - RITA1_peaks
 - RITA_1_6-8h_
 - p110
 - Tracks
 - Journal
 - 2010.11.03_21:22:1
 - 2010.11.04_07:17:3
 - 2010.11.04_07:26:3
 - 2010.11.04_07:28:0

Tables: RITA_1_6-8h_vs_control_dn_AND_p53_peak_TF_DT_200_activity Tables: PASS_molecule2activity_cancer_targets RITA: Drug_target
 Tables: RITA_1_6-8h_vs_control_dn_AND_p53_peak_TF_DT RITA: 4DT RITA: p110 RITA: 3DT Tables: RITA_1_6-8h_vs_control_dn_AND_p53_peak_TF_DT_200

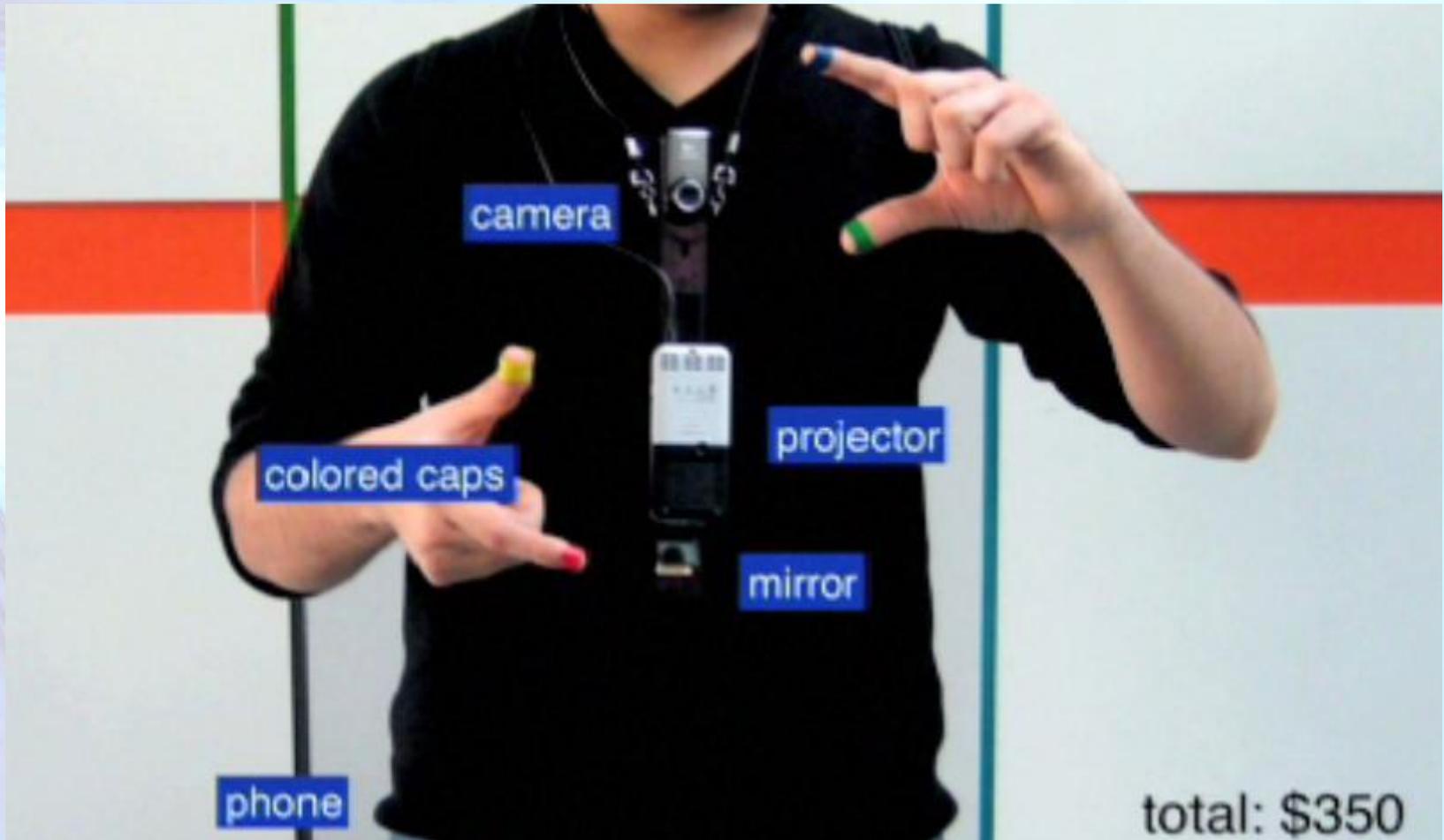


ID: 3DT
Title: 3DT
Complete name: data/Collaboration/RITA/Data/Tables/3DT
Attributes:

- name: 3DT
- node-image:

INFO : Generating result...
 INFO : Drug Target search started
 INFO : 174 molecules from input list are taken for the analysis
 INFO : 1874 regulator node candidates found
 INFO : Calculating scores...
 INFO : DataCollection RITA_1_6-8h_vs_control_dn_AND_p53_peak_TF_DT created

Augmented Reality



Processing Packaging Information



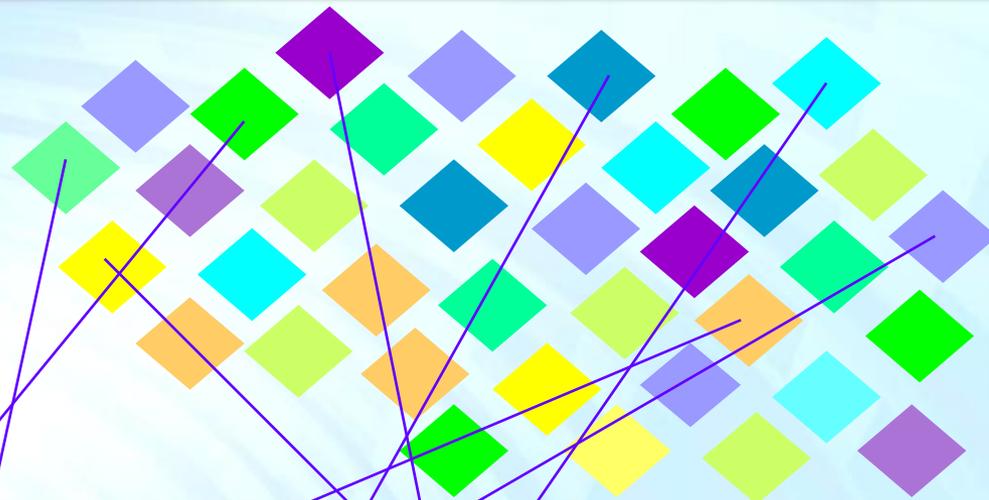
MIT Media Lab

Creation of VO from Collaboration Pool

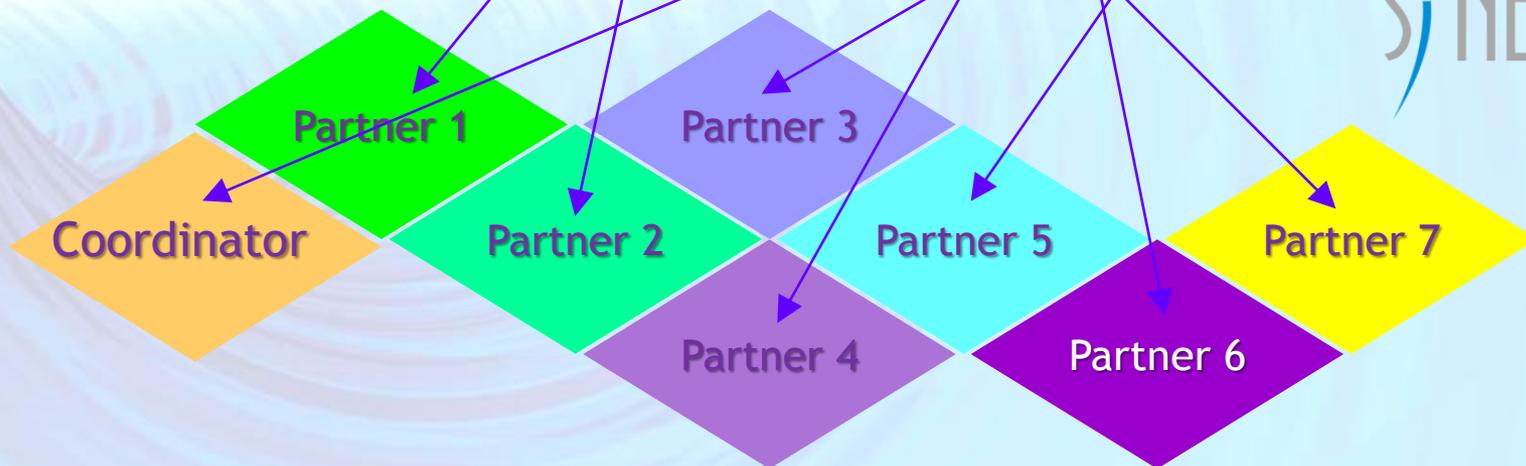
Network

Opportunity

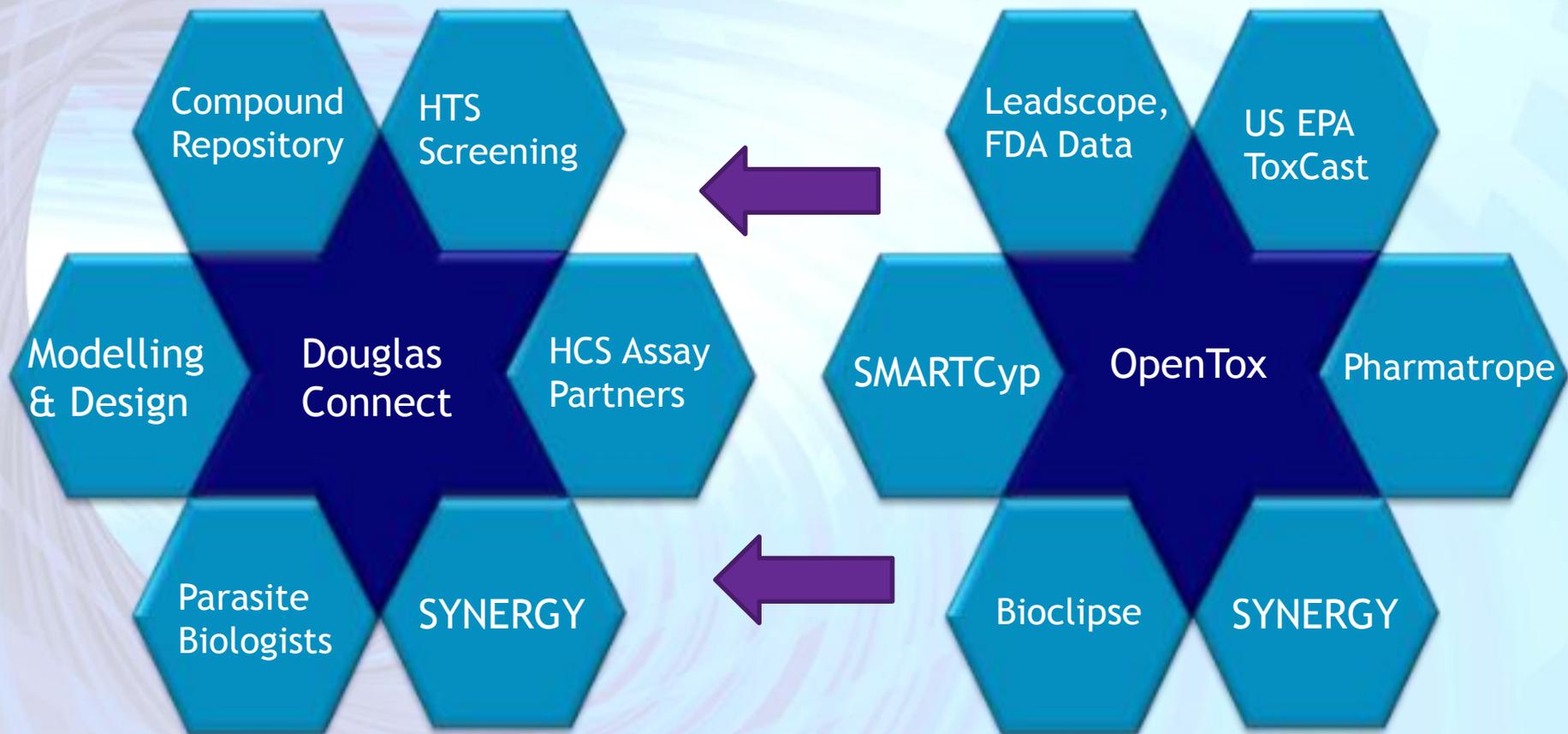
Call for Tender
Need for joint effort
Major project



Virtual Organisation



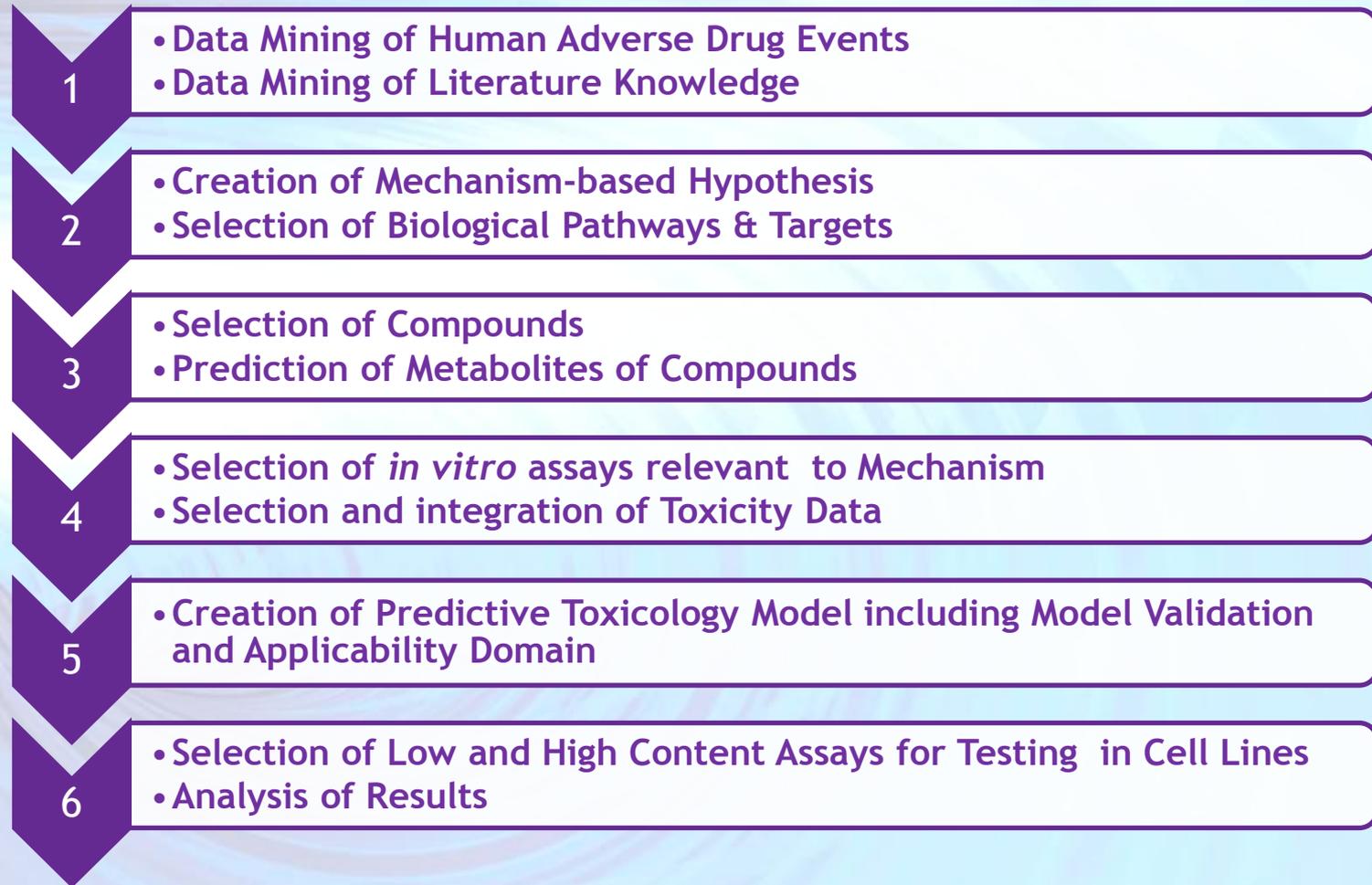
Virtual Organisation Pilots



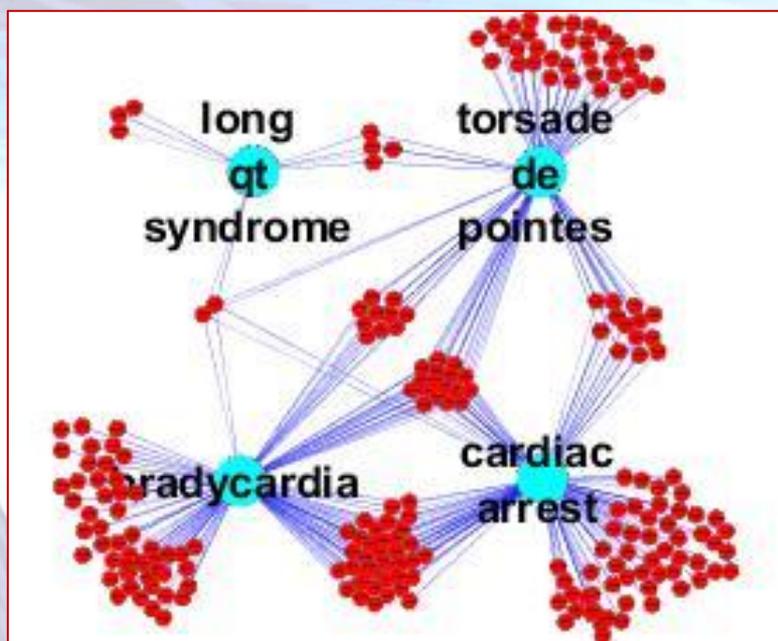
Neglected Disease Drug Design VO

Predictive Toxicology VO

OpenTox - Synergy Predictive Toxicology VO Pilot Strategy Development & Case Study



Analysis of Adverse Events Based on Pharmacological Activity

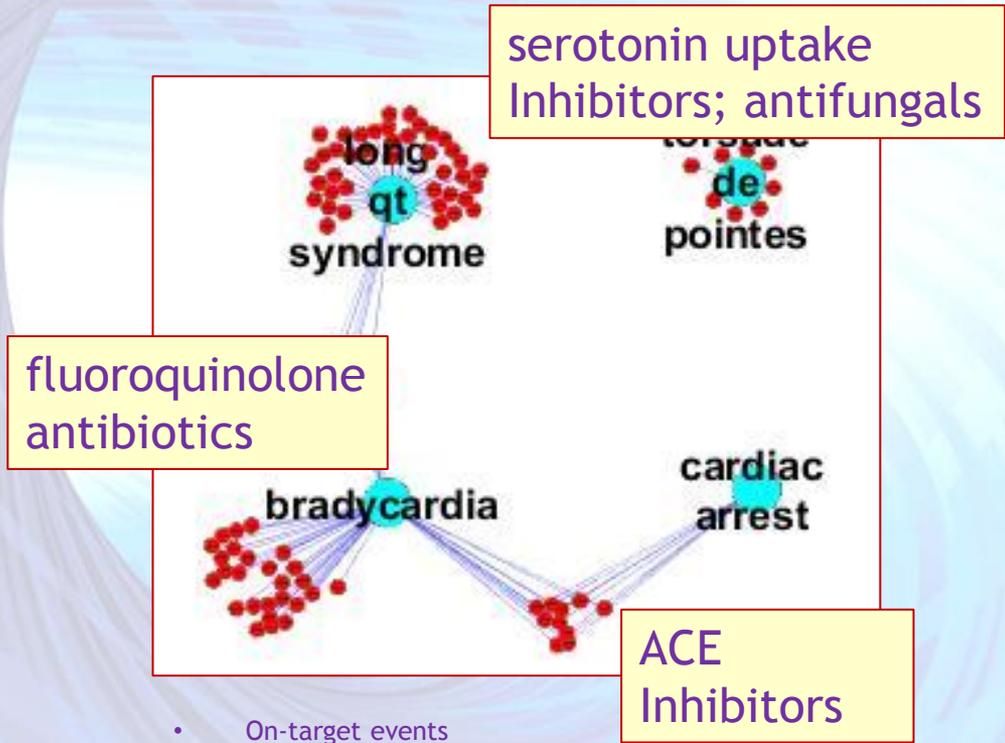


- Question addressed:
 - Are the adverse events a function of inhibiting the pharmacological target
 - Or is the adverse event due to an off-target activity

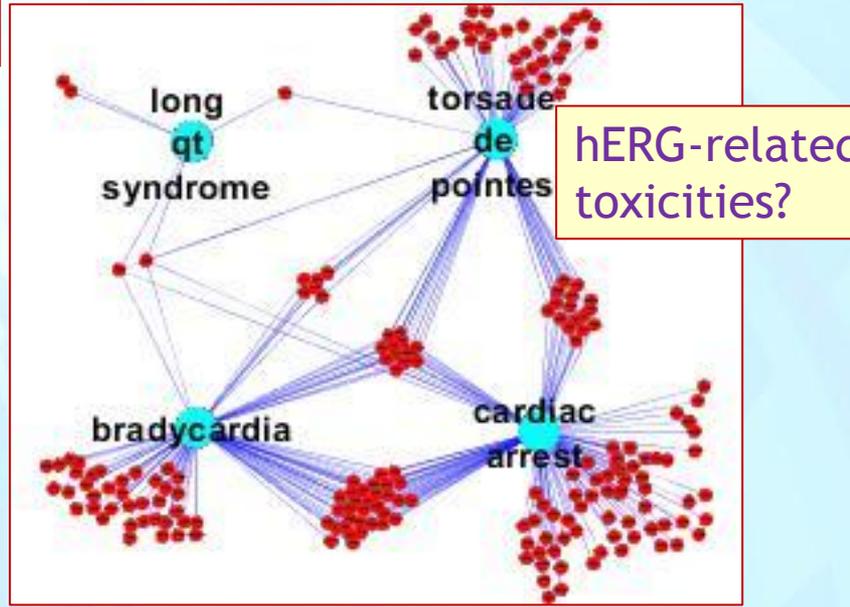
- Cardiac adverse events
- Related to hERG ion channel?

cyan = adverse event, red = drug
lines define links

Example: Cardiac Adverse Events



cyan = adverse event, red = drug
lines define links



Ontology Issues

- MeSH terms inadequate to describe drug activity
 - Higher granularity definitions required
 - e.g. “ACE inhibitor” instead of “antihypertensive agent”
 - Are alternative definitions available?
- MedDRA terms are used to define adverse events
 - Drug pharmacological action needs to be defined in the same terms

SAM VO targeting Plasmodium Kinases

a Sporozoites

Sporozoite infectivity
▶ **PbCDPK6^{coppi}**

Liver

Merozoites

Trophozoite

Game

Erythrocytic schizogony:

- ▶ Pfmap-2
- ▶ PfCDPK1^{Winzeler}
- ▶ PfPK7
- ▶ Pbcrk-1
- ▶ Pfnek-1
- ▶ PfPK5
- ▶ PfPK6
- ▶ PfPK9
- ▶ PfCK2
- ▶ PfGSK3
- ▶ PfTKL1
- ▶ PfTKL3
- ▶ Pfcrk-3
- ▶ Pfcrk-4
- ▶ PfARK1
- ▶ PfARK2
- ▶ PfARK4
- ▶ PfPK4

Ookinete maturation:

- ▶ **Pbnek-4**
- ▶ **Pbnek-2**

Gametogenesis:

- ▶ **PbDCPK4^{Billker}**
- ▶ **Pbmap-2**
- ▶ PfPKG^{Baker}

b

Salivary glands

Non-essential for erythrocytic schizogony

- ▶ Pfmap-1
- ▶ PfPK7
- ▶ **PbCDPK3^{Ishino, Billker}**
- ▶ **Pfnek-4 / Pbnek-4**
- ▶ **PbDCPK4^{Billker}**
- ▶ **Pbmap-2**
- ▶ PfPKG^{Baker}
- ▶ Pfnek-2
- ▶ Pfnek-3
- ▶ Pfnek-4
- ▶ Pfcrk-5
- ▶ PfeIK1
- ▶ PfeIK2
- ▶ PfTKL-2
- ▶ PfTKL-4
- ▶ PfTKL-5

Oocyst maturation:

- ▶ PfPK7

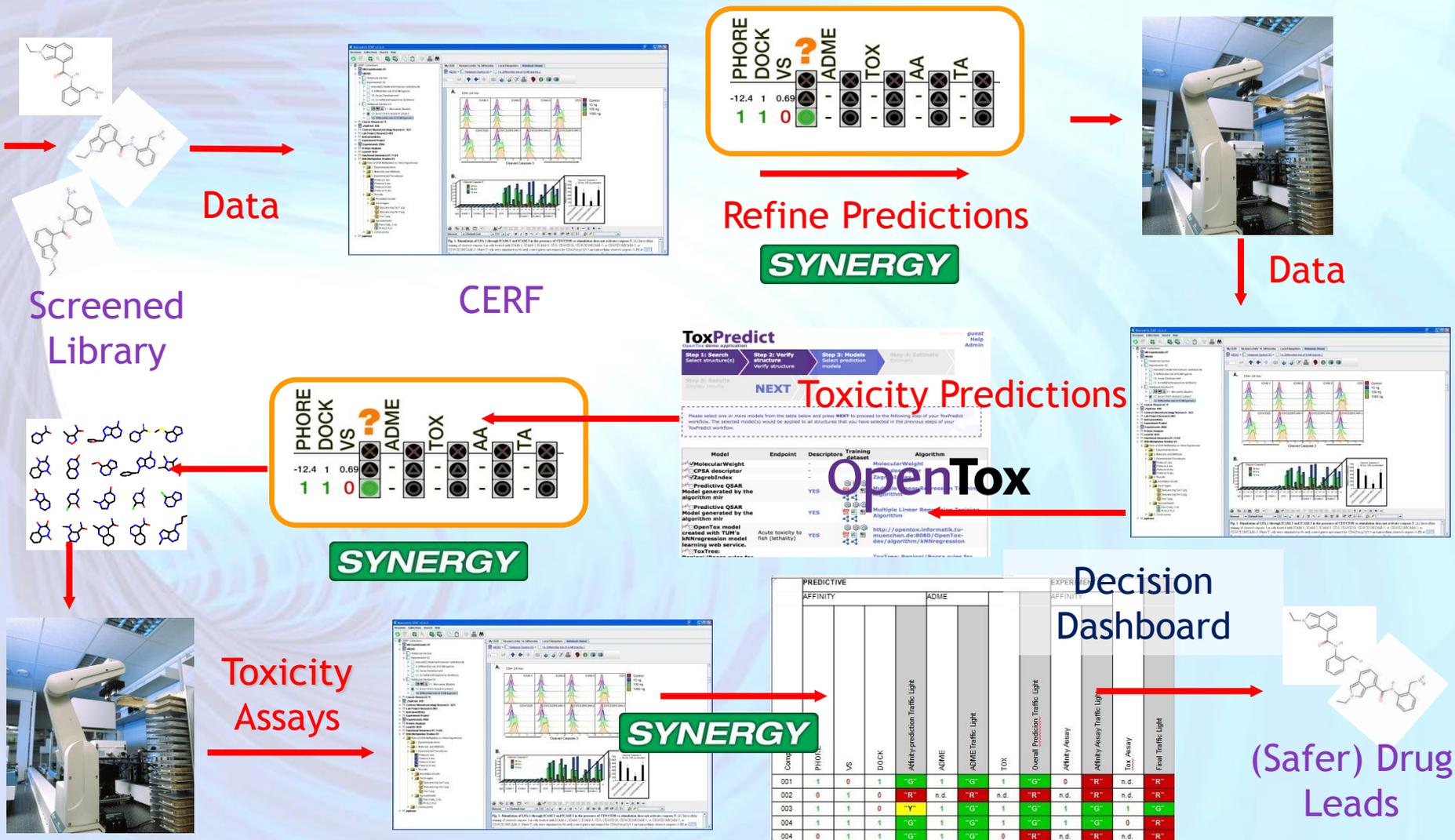
Ookinete migration:

- ▶ **PbCDPK3^{Ishino, Billker}**

Zygote

Gametes

Synergy Drug Design Collaboration Pilot



Recording of Collaborative R&D

Controlled Vocabularies

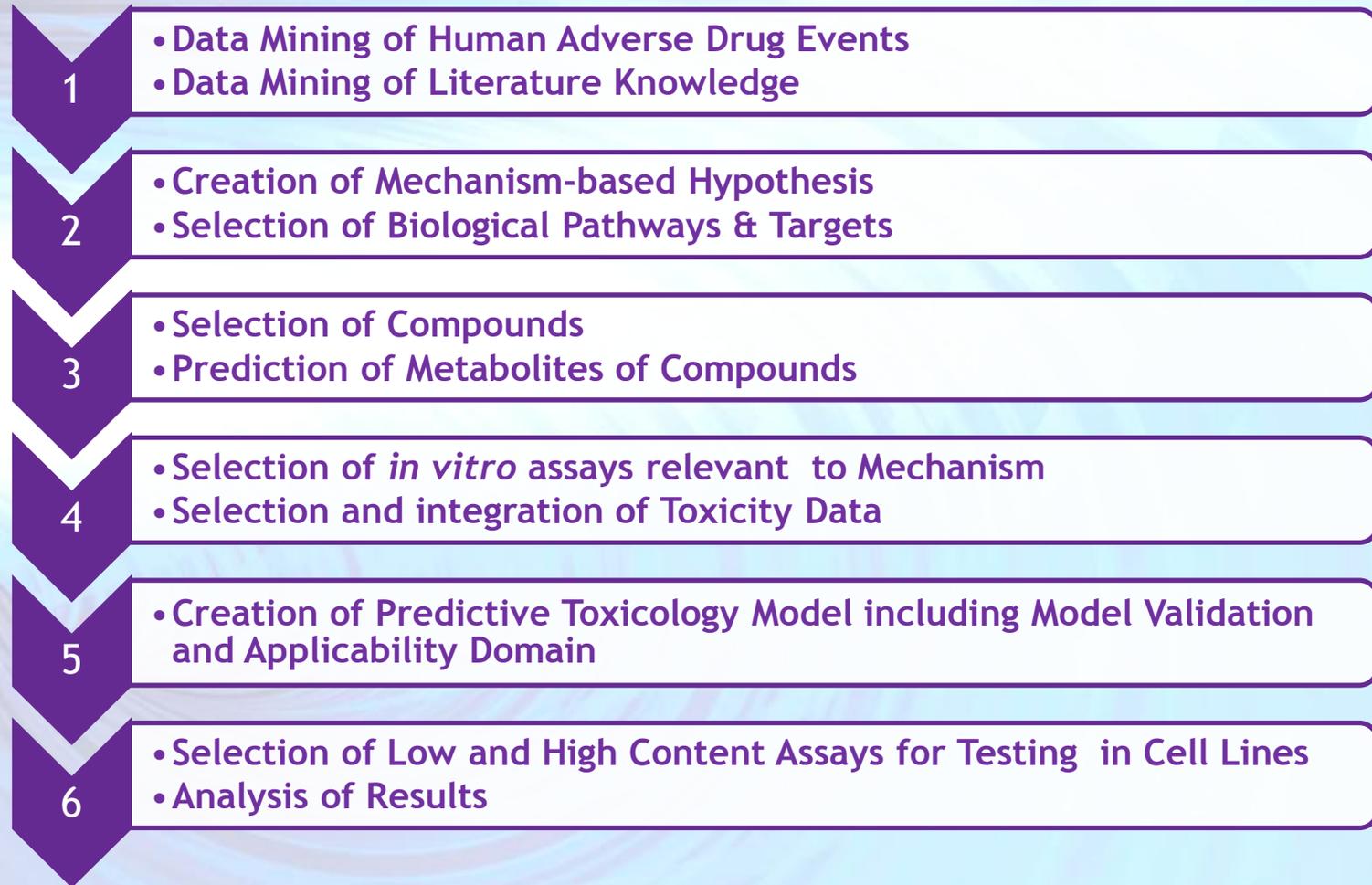
Visualisation

Collaborative Electronic Laboratory Notebook (ELN)

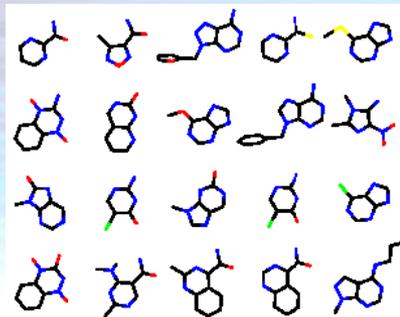
Fig. 1. Stimulation of LFA-1 through ICAM-2 and ICAM-3 in the presence of CD3/CD28 co-stimulation does not activate caspase-3. (A) Intracellular staining of cleaved-caspase 3 in cells treated with ICAM-1, ICAM-2, ICAM-3, CD3, CD3/CD28, CD3/CD28/ICAM-1, or CD3/CD28/ICAM-2, or CD3/CD28/ICAM-3. Naive T cells were stimulated in 96-well coated plates and stained for CD4-PerCPy5.5 and intracellular cleaved-caspase-3-FE at 24 hrs.

Condition	24 hrs	48 hrs	72 hrs
IgG	~100	~100	~100
ICAM-1	~100	~100	~100
ICAM-2	~100	~100	~100
ICAM-3	~100	~100	~100
CD3	~100	~100	~100
CD3/CD28	~100	~100	~100
CD3/CD28/ICAM-1	~100	~100	~100
CD3/CD28/ICAM-2	~100	~100	~100
CD3/CD28/ICAM-3	~100	~100	~100

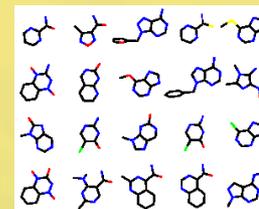
OpenTox - Synergy Predictive Toxicology VO Pilot Strategy Development & Case Study



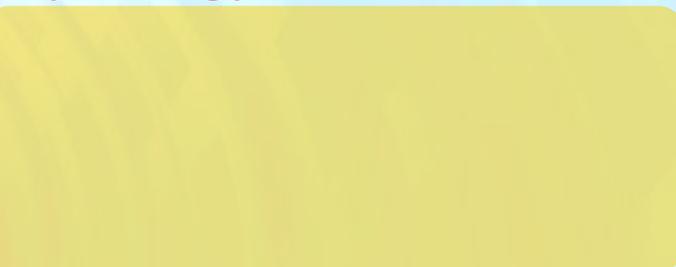
1. A library of compounds is entered to the ELN



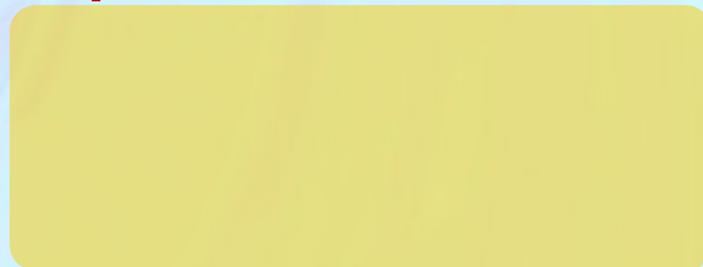
ELN



Synergy

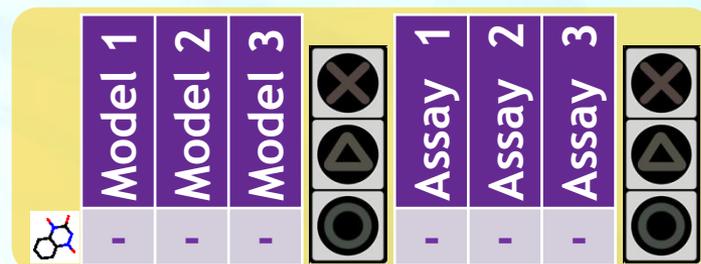


OpenTox

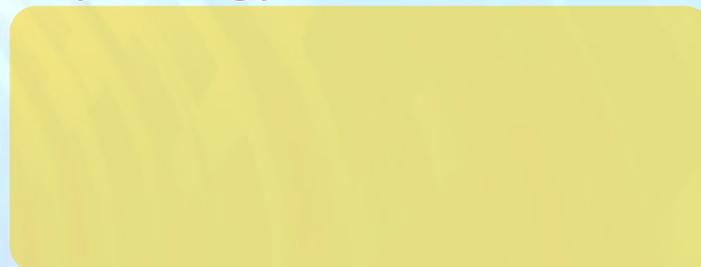


2. Each compound is assigned a data structure in ELN

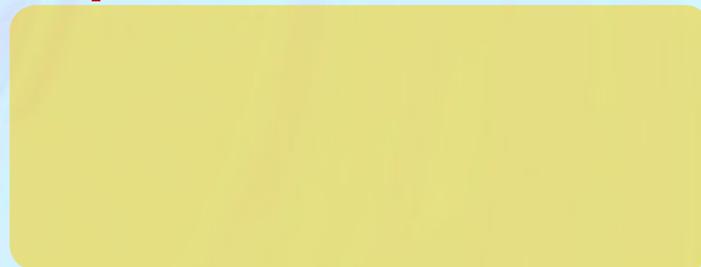
ELN



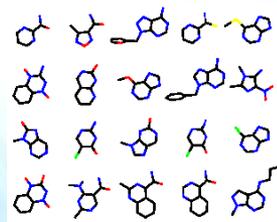
Synergy



OpenTox



3. ELN passes compounds to OpenTox and SYNERGY



ELN

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

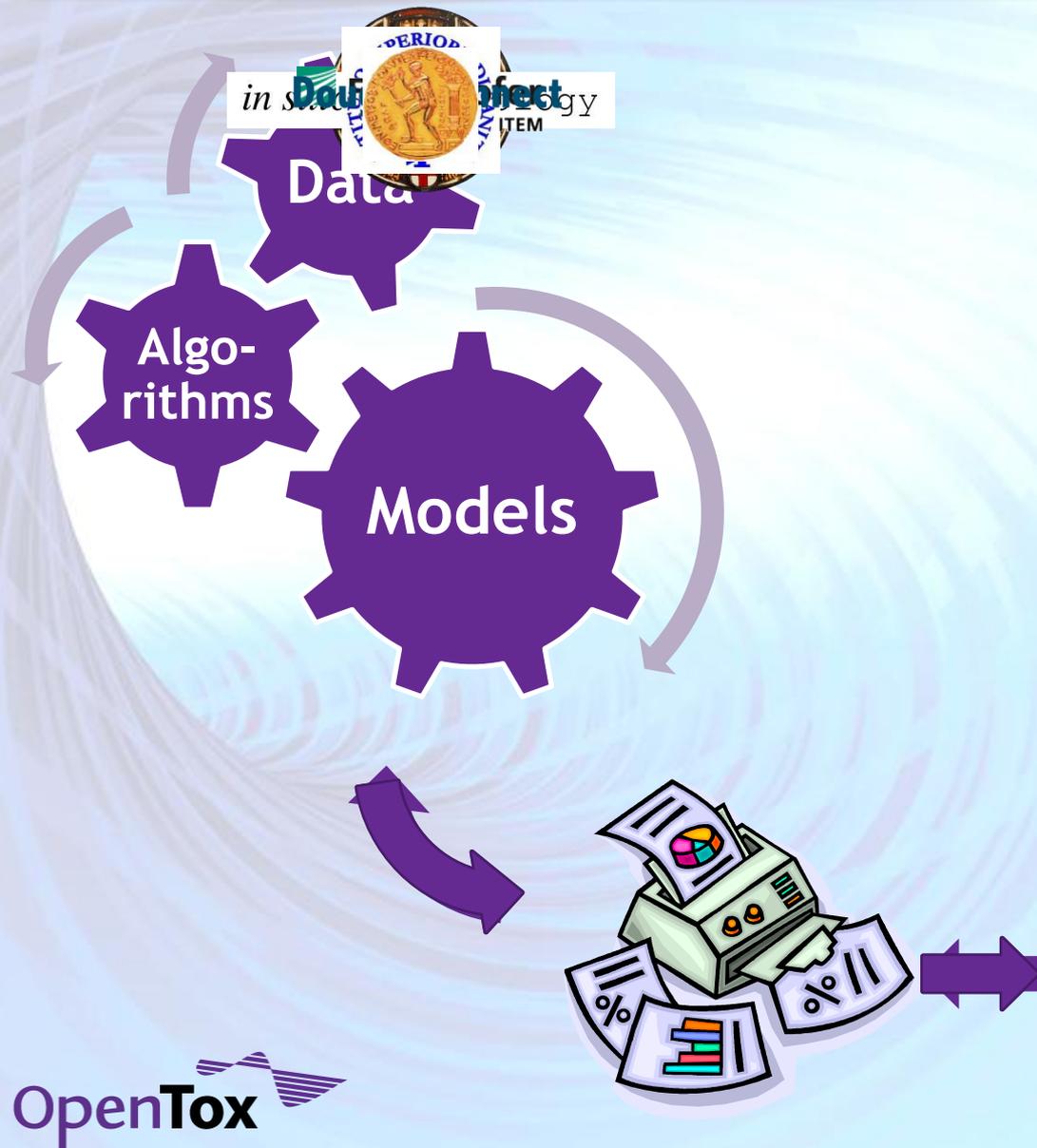
Synergy

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

OpenTox

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

4. OpenTox computes toxicity predictions



ELN

	Model 1	Model 2	Model 3	<input type="checkbox"/>	Assay 1	Assay 2	Assay 3	<input type="checkbox"/>
	-	-	-	<input type="radio"/>	-	-	-	<input type="radio"/>

Synergy

	Model 1	Model 2	Model 3	<input type="checkbox"/>	Assay 1	Assay 2	Assay 3	<input type="checkbox"/>
	-	-	-	<input type="radio"/>	-	-	-	<input type="radio"/>

OpenTox

	Model 1	Model 2	Model 3	<input type="checkbox"/>	Assay 1	Assay 2	Assay 3	<input type="checkbox"/>
	1	0	1	<input type="radio"/>	-	-	-	<input type="radio"/>

5. OpenTox sends back a report to ELN



ELN

	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		
	-	-	-			-	-	-		

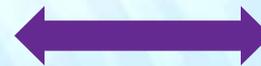
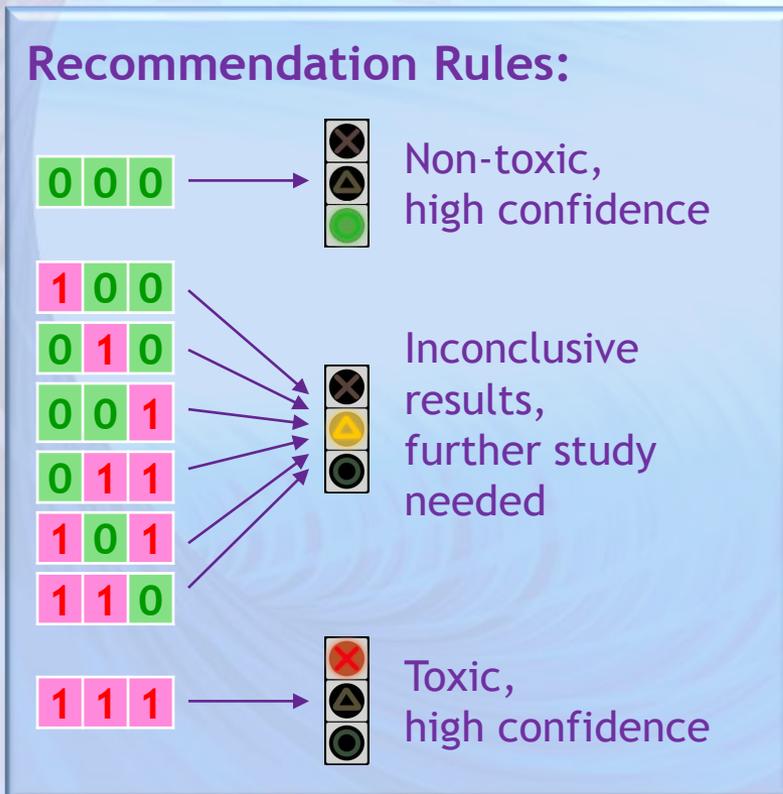
OpenTox

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

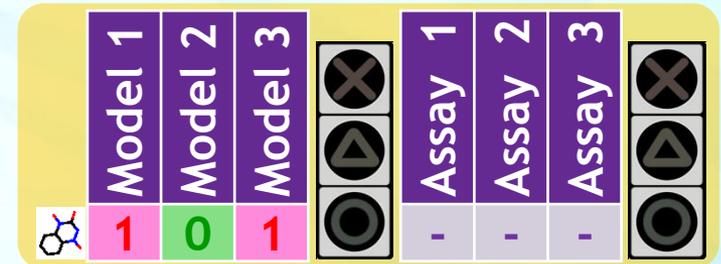
6. ELN sends the results to SYNERGY



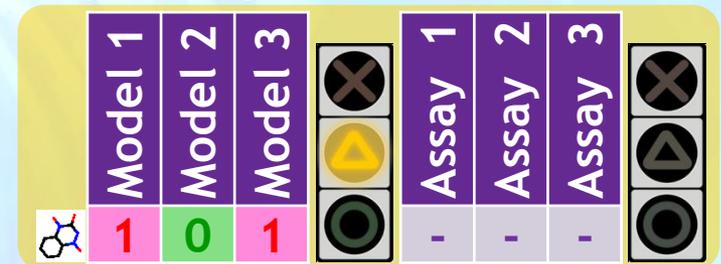
7. SYNERGY applies the Recommendation Rules



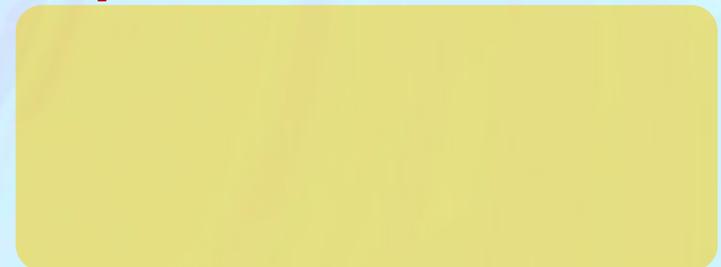
ELN



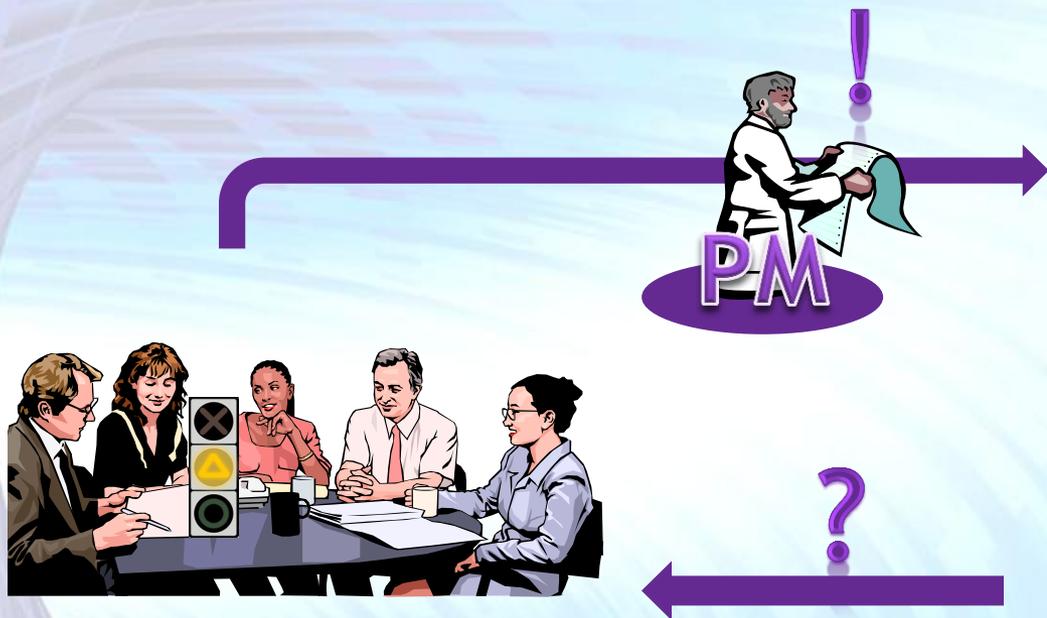
Synergy



OpenTox



8. Inconclusive data → SYNERGY calls a meeting



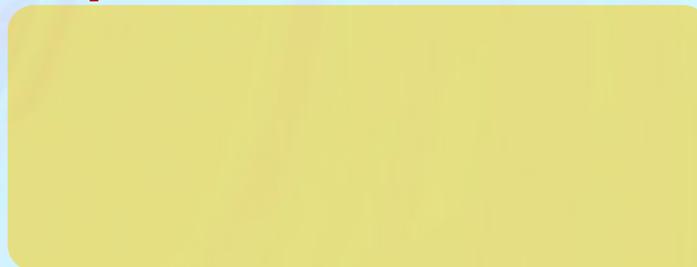
ELN

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

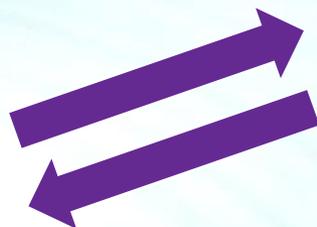
Synergy

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

OpenTox



9. Experimental assays confirm toxicity



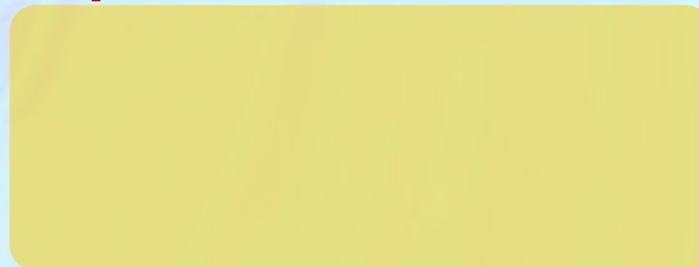
ELN

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

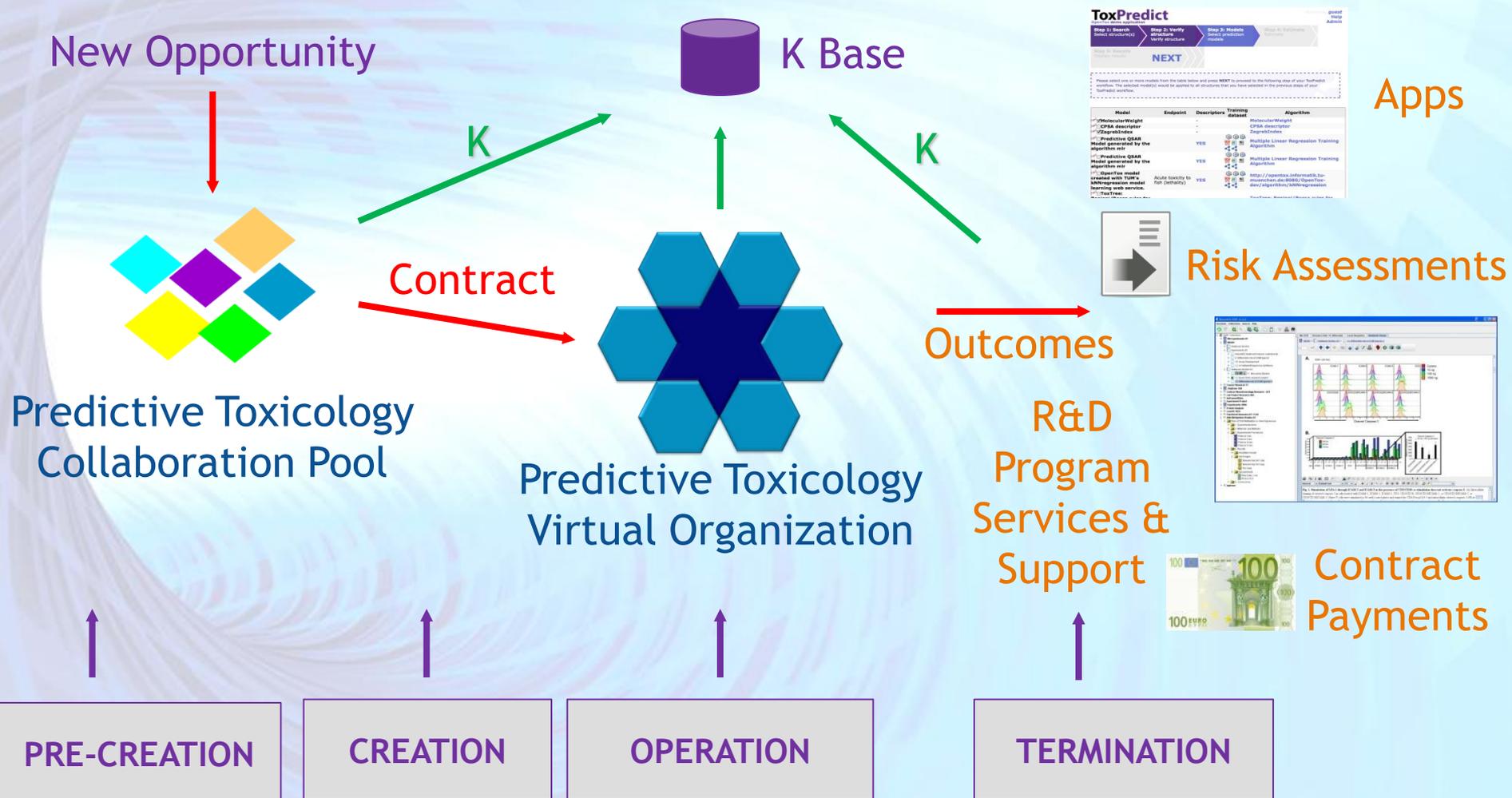
Synergy

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

OpenTox



Sustainability Model



ToxPredict

Step 1: Search
Select structure(s)

Step 2: Verify structure
Verify structure

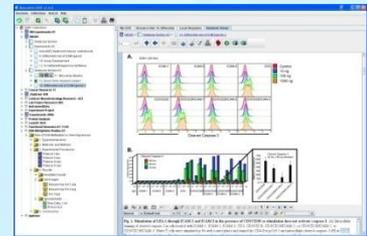
Step 3: Train
Train prediction model

Step 4: Evaluate

NEXT

Please click one or more models from the table below and click **NEXT** to proceed to the following step of your "Virtual workflow". The selected model(s) will be added to the structure that you have selected in the previous steps of your "Virtual workflow".

Model	Endpoint	Descriptors	Training dataset	Algorithm
<input type="checkbox"/> Predictive Weight	-	-	RobertMarrugo	Multiple Linear Regression Training Algorithm
<input type="checkbox"/> CPBA descriptor	-	-	CPBA descriptor	Multiple Linear Regression Training Algorithm
<input type="checkbox"/> ClogP descriptor	-	-	CPBA descriptor	Multiple Linear Regression Training Algorithm
<input type="checkbox"/> Predictive QSAR Model generated by the algorithm mlr	YES	CPBA descriptor	Multiple Linear Regression Training Algorithm	Multiple Linear Regression Training Algorithm
<input type="checkbox"/> Predictive QSAR Model generated by the algorithm mlr	YES	CPBA descriptor	Multiple Linear Regression Training Algorithm	Multiple Linear Regression Training Algorithm
<input type="checkbox"/> QSAR model created with TSM's Adaptive model learning web service	YES	Acute toxicity to fish (peracety)	https://apertea.informatica.uniroma3.it/0000/Operativ.../algorithms/AdaptiveRegression	Adaptive Regression
<input type="checkbox"/> DTSTree	-	-	-	-



Our Drivers - Taking on Technical, Cultural and “Other” Challenges of the Unexpected



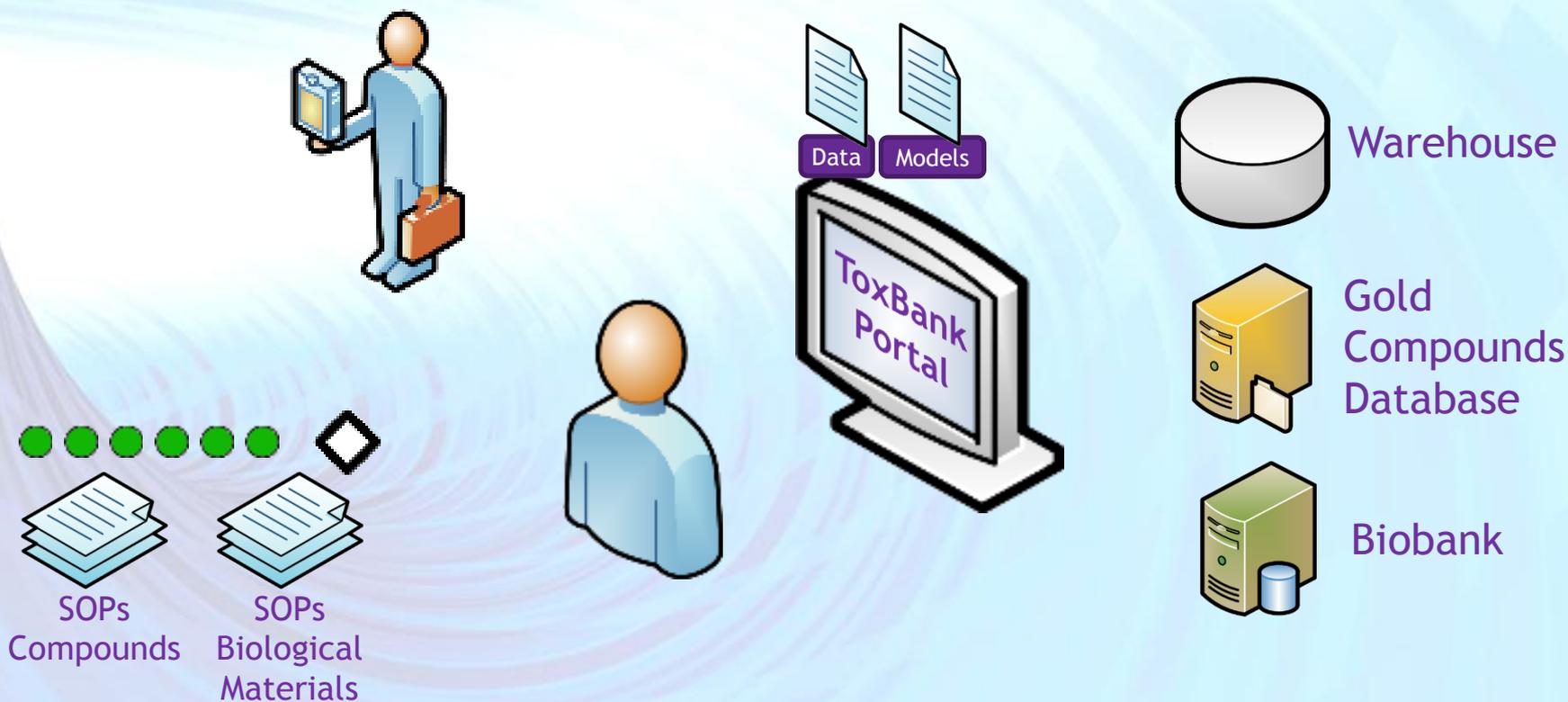
[Visit with Lions at Mukuni Reintroduction Project, Livingstone, Zambia](#)

Our Drivers - Taking on Technical, Cultural and “Other” Challenges of the Unexpected



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

User orders compounds and biological materials



ToxBank Infrastructure Project (scheduled for a Jan 2011 start)

Establishment of a ...



- ... cell and tissue banking information resource
- ... repository for the selected test compounds
- ... database of reference test compounds
- ... dedicated web-based data warehouse

Our Funding Support...

For more information, visit

www.opentox.org

Contact me:

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

twitter.com/barryhardy

barryhardy.blogs.com



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).