

Development and Use of Predictive Toxicology Applications

OpenTox Presentation

19 October 2010

University Uppsala, Sweden

Barry Hardy (Douglas Connect)

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

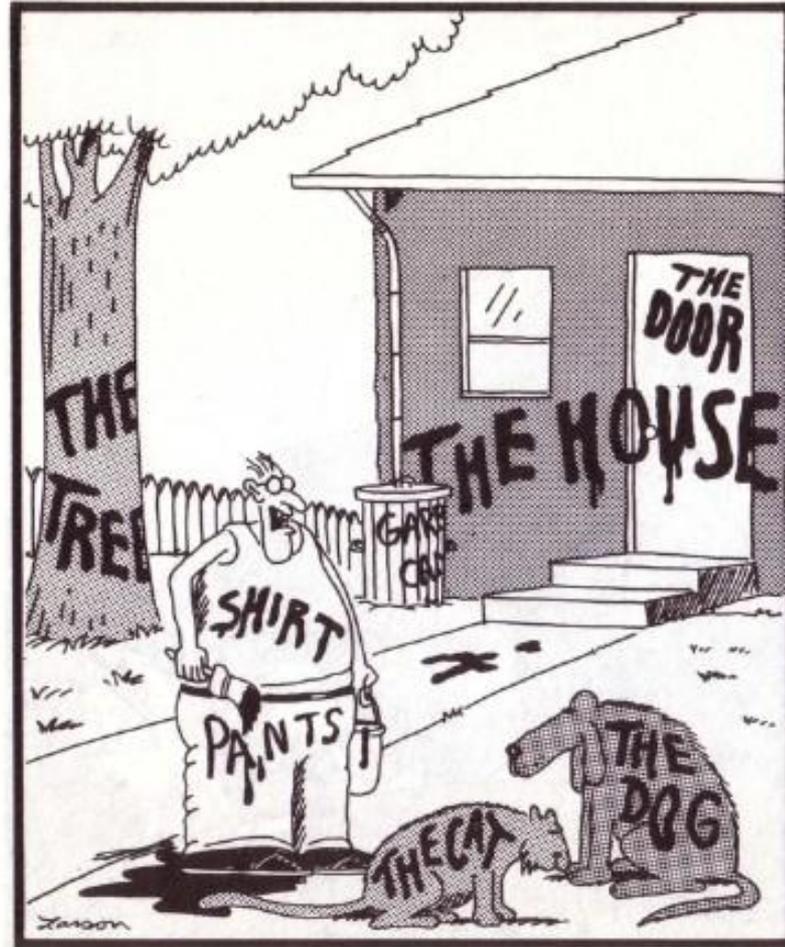
21742.0
Benzene
benzene
OK

MolecularWeight  **MolecularWeight**

MW

78.1112

Semantic Reflections



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

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

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

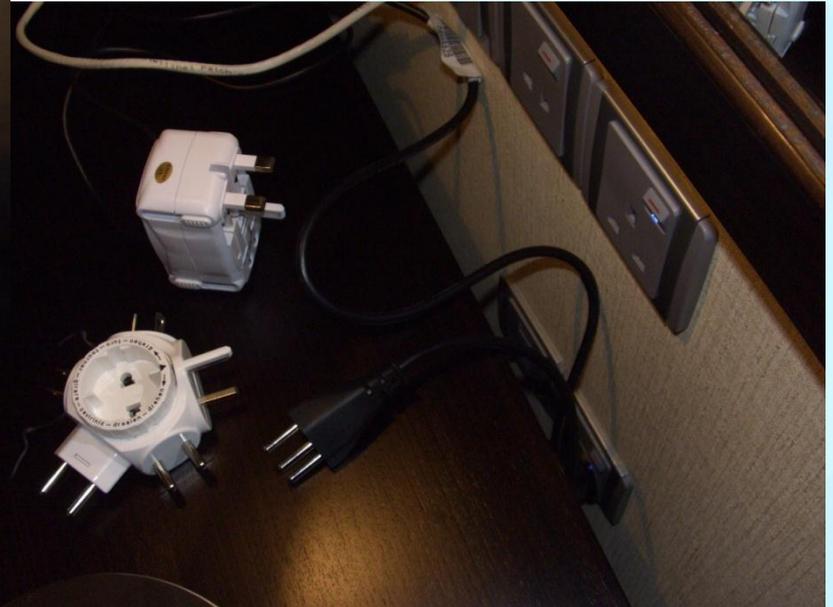
Journal of Cheminformatics Publication

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

Absence of Interoperability creates Problems



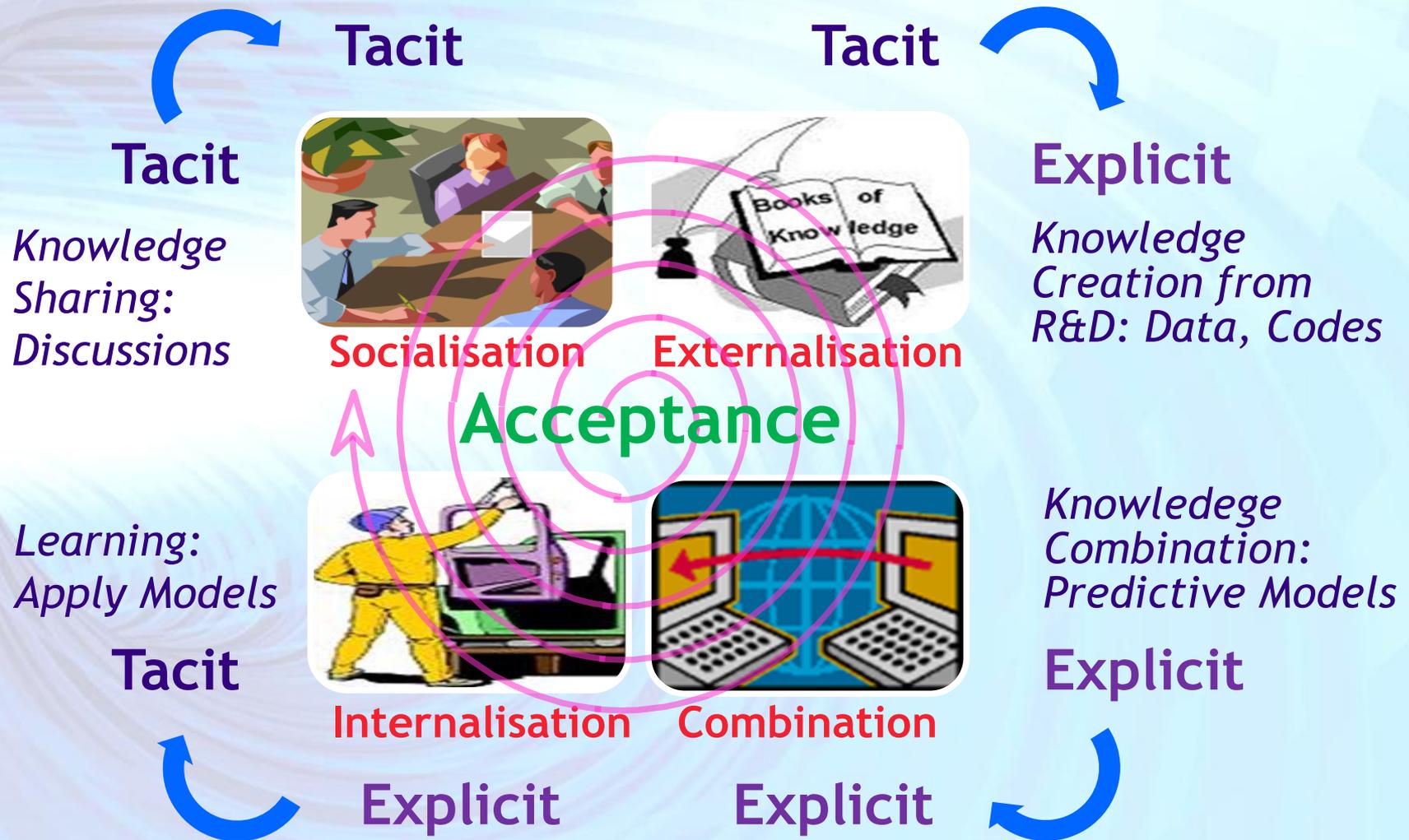
Adaptor Challenge in Jeddah, 2008

Interacting Components create Solutions



Adaptor Solution in Jeddah, 2008

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



Collaborative Predictive Toxicology Challenge

Input Structure



Out - Toxic or Not?

- LD50
- Liver Toxicity
- Secondary Metabolites
- Bioavailability
- Mutagenicity
- Carcogenicity
- Reproductive Toxicology
- Skin Irritation
- Aqua Toxicity
- Combined predictions for arrays of multiple end points



Driver



Increasing demands on industry to satisfy safety evaluation and risk assessment required by REACH legislation. (Over 140k cmpds registered).

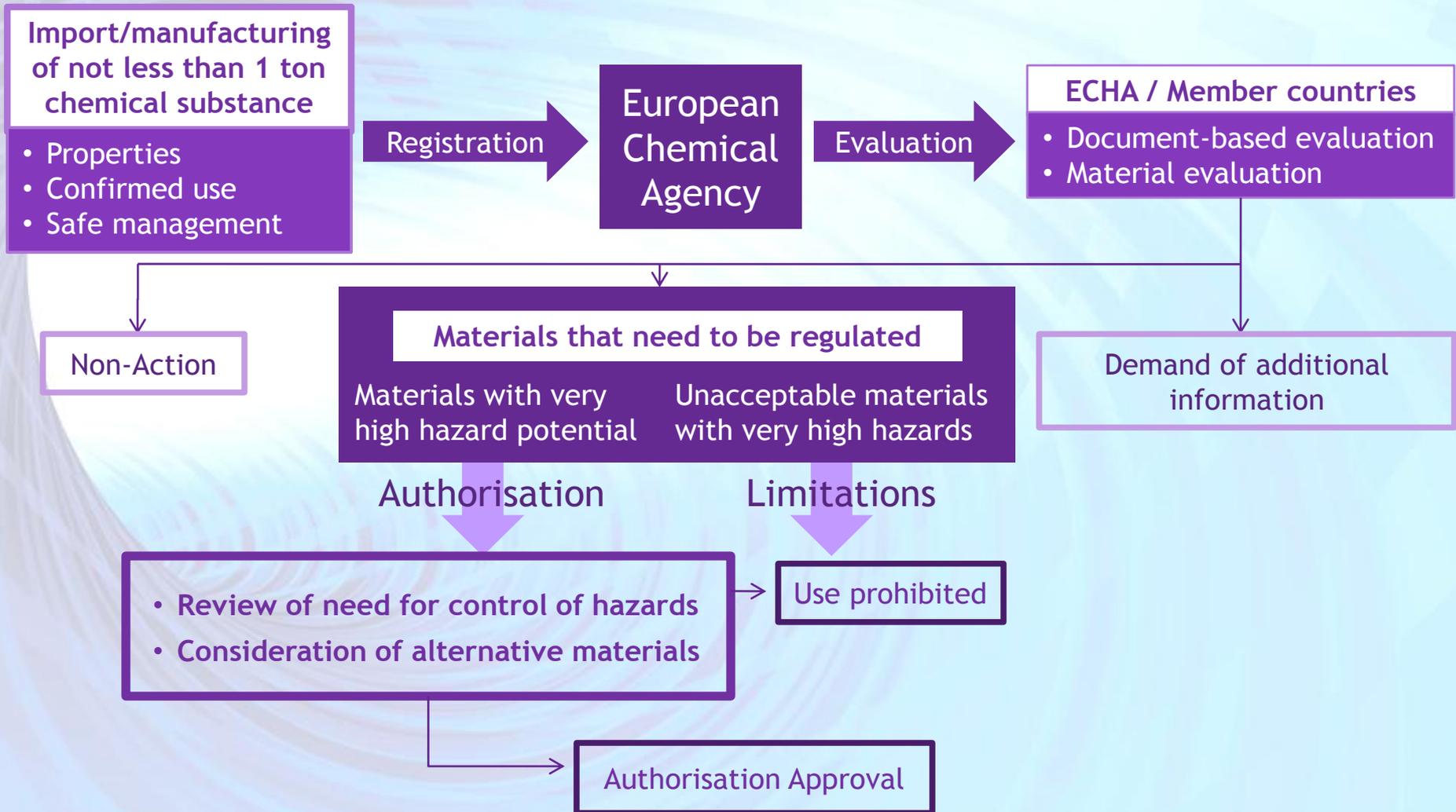
REACH

An aerial photograph of a lush green landscape, possibly a park or a rural area, viewed through a black grid pattern. The word "REACH" is written in large, white, sans-serif capital letters across the center of the image.

REACH

(enRegistrement, Evaluation et
Autorisation des substances CHimiques)

REACH Registration



Challenges to Integrated Resources & Applications

- Database silos
- Missing information
- Varying quality
- Hard to integrate data
- Hard to integrate models
- No common framework
- Lack of standards
- Lack of validation
- Complex subject
- Application difficult
- Lack of transparency
- Interdisciplinary collaboration

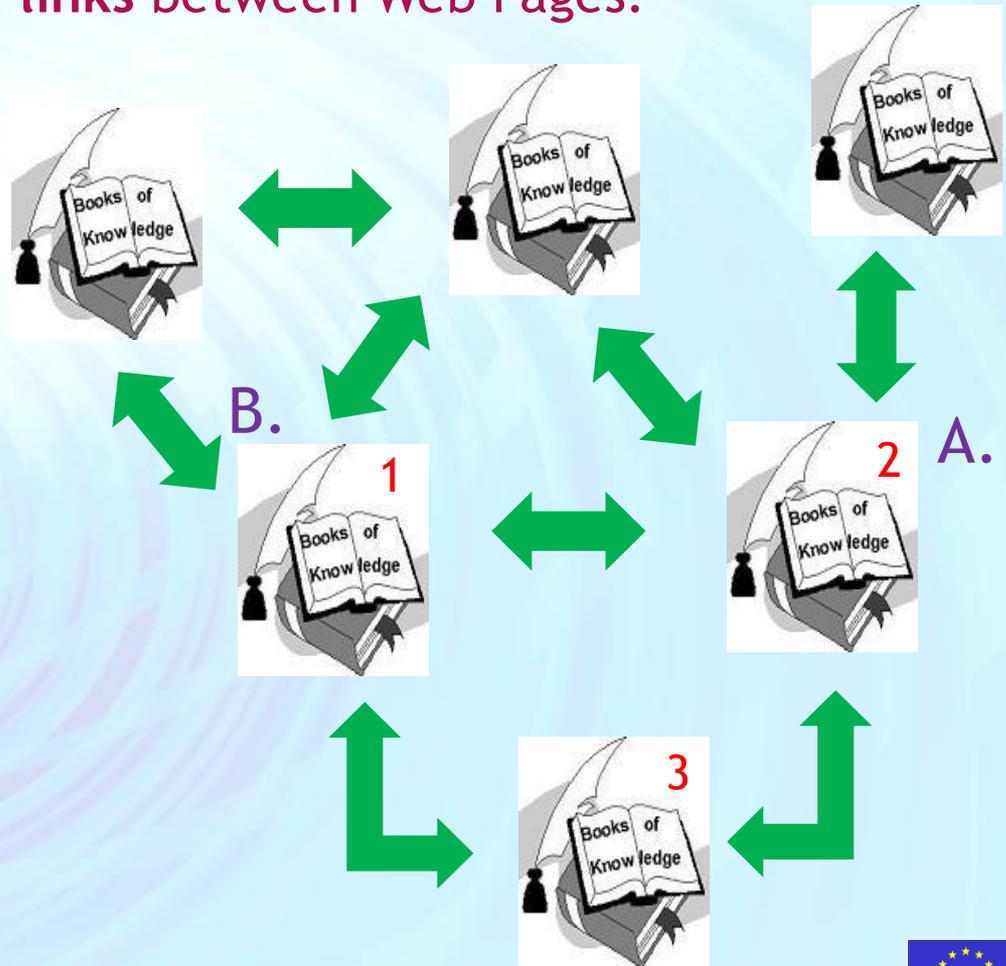
Value is in Linking

The key idea of Google's founders in creating their search engine:
There is useful knowledge in the **links** between Web Pages.

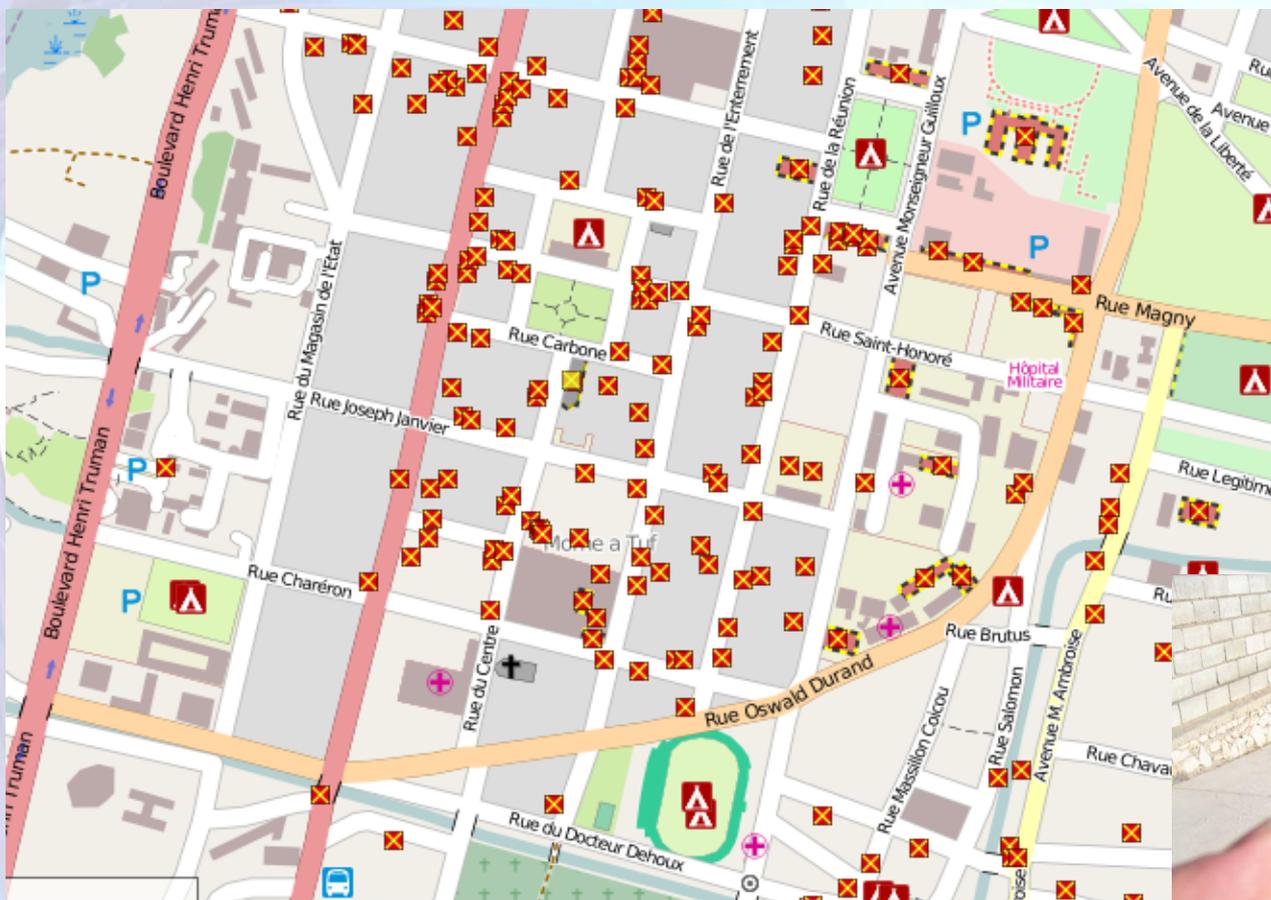
Page Ranking

A page is ranked higher in a search if:

- A. it has more connections to it than other pages
- B. the pages connecting to it have higher ranking themselves



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

- Toxicologist, Biologist, 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

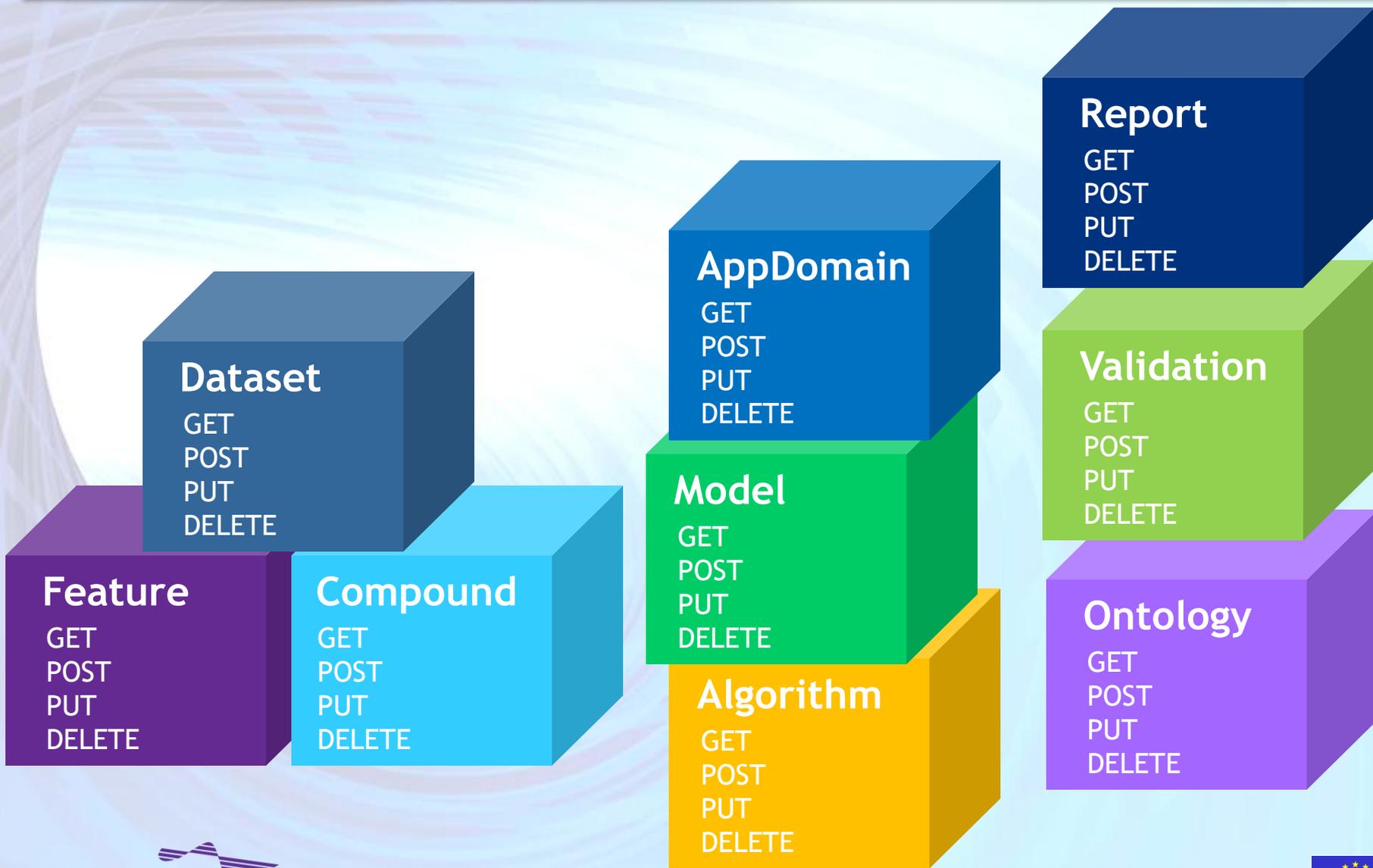
Validation: Methods for estimating the accuracy of model predictions

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

Interface Definitions

Description	Method	URI	Parameter	Result	Status codes
Get available feature URIs for a compound	GET	/compound/{cid}/feature	?feature_uris[]="URI to features" (optional)	Returns representation of the features as uri-list or RDF All available features are returned, if no parameter is specified.	200,404,503
Create a new feature value	POST	/compound/{cid}/feature	?feature_uri="URI to feature" (mandatory, single feature)&value=the_value	URI of the compound with the new feature, e.g. /compound/{id}?feature_uris[]=the-new-feature	200,400,503
Update a new feature value	PUT	/compound/{cid}/feature	?feature_uri="URI to feature" (mandatory, single feature)&value=the_value		200,400,404,503
Delete specified features from the compound	DELETE	/compound/{cid}/feature	?feature_uris[]="URI to features" (optional)		200,400,404,503

Interface Definitions

Description	Method	URI	Parameters	Result	Status codes
get description of a specific feature definition	GET	/feature/{id}	-	URI-list or RDF representation of a feature.	200,404,503
create a new feature	POST	/feature	Content-type ="any-of-RDF-types", content=RDF-representation	URI of the new feature definition.	200,400,404,503
update feature	PUT	/feature/{id}	Content-type ="any-of-RDF-types", content=RDF-representation	-	200,400,404,503
delete feature	DELETE	/feature/{id}	-	-	200,400,404,503
get a list of available feature definitions	GET	/feature	? query =URI-of-the-owl:sameAs-entry	URI list or RDF of features found by the query or all available, if query is empty. Returns all features, for which owl:sameAs is given by the query.	200,404,503

Interface Definitions

Description	Method	URI	Parameters	Result	Status codes
Get a list of available datasets	GET	/dataset	Query parameters (optional, to be defined by service providers).	List of URIs or RDF for the metadata only.	200,404,503
Get a dataset	GET	/dataset/{id}	-	Representation of the dataset in a supported MIME type.	200,404,503
Query a dataset	GET	/dataset/{id}	compound_uris[] and/or feature_uris[] to select compounds and features; further query parameters may be defined by service providers.	Representation of the query result in a supported MIME type.	200,404,503
Get metadata for a dataset	GET	/dataset/{id}/metadata	-	Representation of the dataset metadata in a supported MIME type.	200,404,503
Get a list of all compounds in a dataset	GET	/dataset/{id}/compounds	-	List of compound URIs.	200,404,503
Get a list of all features in a dataset	GET	/dataset/{id}/features	-	RDF or List of feature URIs (pointing to feature definitions/ontologies).	200,404,503

Interface Definitions

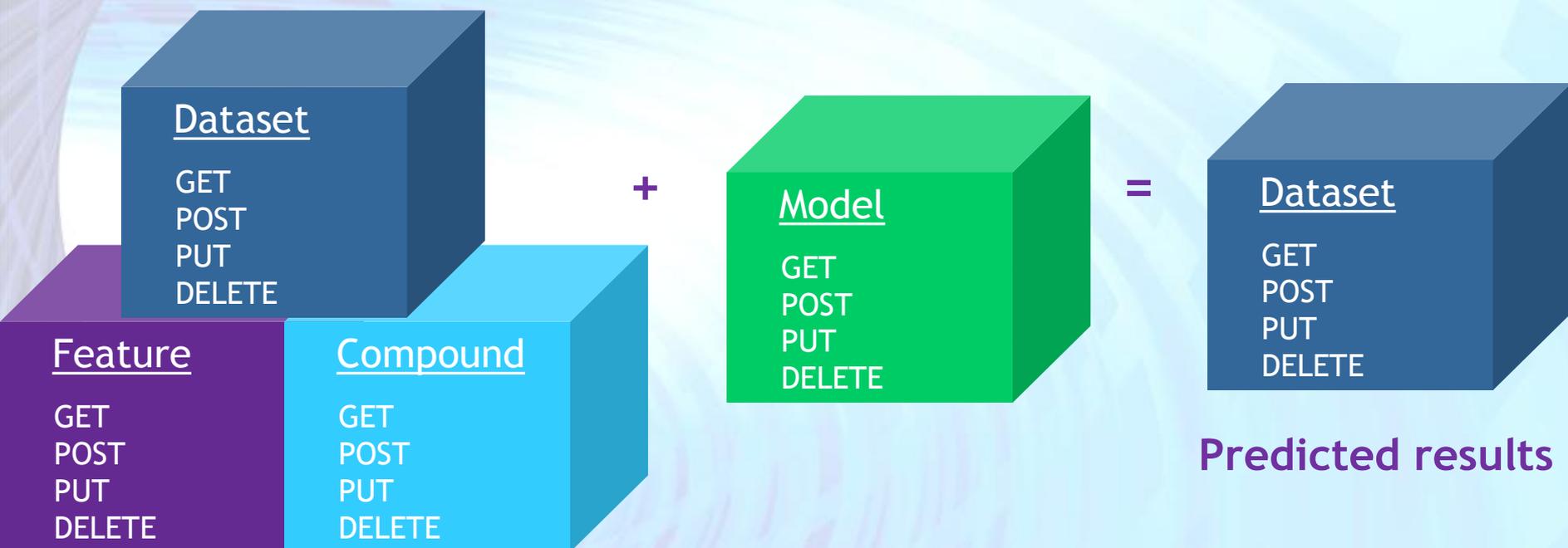
Description	Method	URI	Parameters	Result	Status codes
Get a list of all available models	GET	/model	(optional) ?query=URI-of-the-owl:sameAs-entry	List of model URIs or RDF representation. If query specified, returns all models, for which owl:sameAs is given by the query.	200,404,503
Get the representation of a model	GET	/model/{id}	-	Representation of the model in a supported MIME type.	200,404,503
Delete a model	DELETE	/model/{id}	-	-	200,404,503
Apply a model to predict a dataset	POST	/model/{id}	dataset_uri result_dataset=dataseturi dataset_service=datasetserviceuri	URI of created prediction dataset (predictions are features), task URI for time consuming computations.	200,202,400,404,500,503
Apply a model to predict a compound	POST	/model/{id}	compound_uri	Prediction in a supported MIME type; task URI for time consuming computations.	200,202,400,404,500,503

Interface Definitions

Description	Method	URI	Parameters	Result	Status codes
Retrieve SPARQL query results	GET	/ontology	? query =SPARQL_QUERY (mandatory)	RDF representation of the query results.	200,404,500
Predefined query to retrieve all models	GET	/ontology/models		RDF representation of all models.	
Predefined query to retrieve all endpoints	GET	/ontology/endpoints		RDF representation of all endpoints.	
Predefined query to retrieve all algorithms	GET	/ontology/algorithms		RDF representation of all algorithms.	
Submit SPARQL query and/or OpenTox service URL	POST	/ontology	uri []=URL of a OpenTox RDF resource query =SPARQL_QUERY	RDF representation of the query results, if query is specified. if uri [] is specified, the server retrieves a RDF representation and adds it to the RDF storage, thus making it available for the subsequent queries.	200,404,500,502

Uniform access to calculations

Read data from a web address - process - write to a web address



<http://myhost.com/dataset/newcompounds>

<http://myhost.com/model/predictivemodel1>

<http://myhost.com/dataset/predictedresults1>

Uniform approach to models creation

Read data from a web address - process - write to a web address



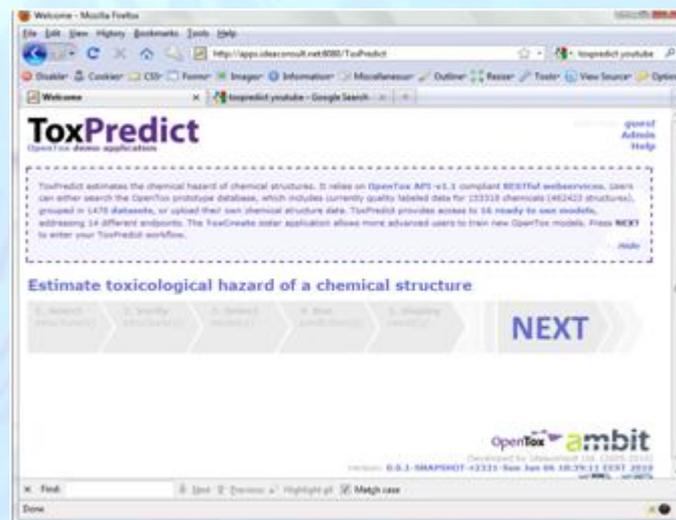
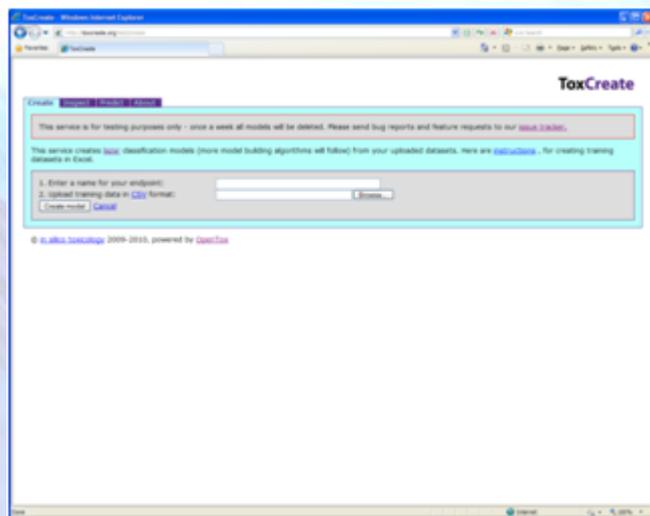
<http://myhost.com/algorithm/neuralnetwork>

<http://myhost.com/dataset/trainingset1>

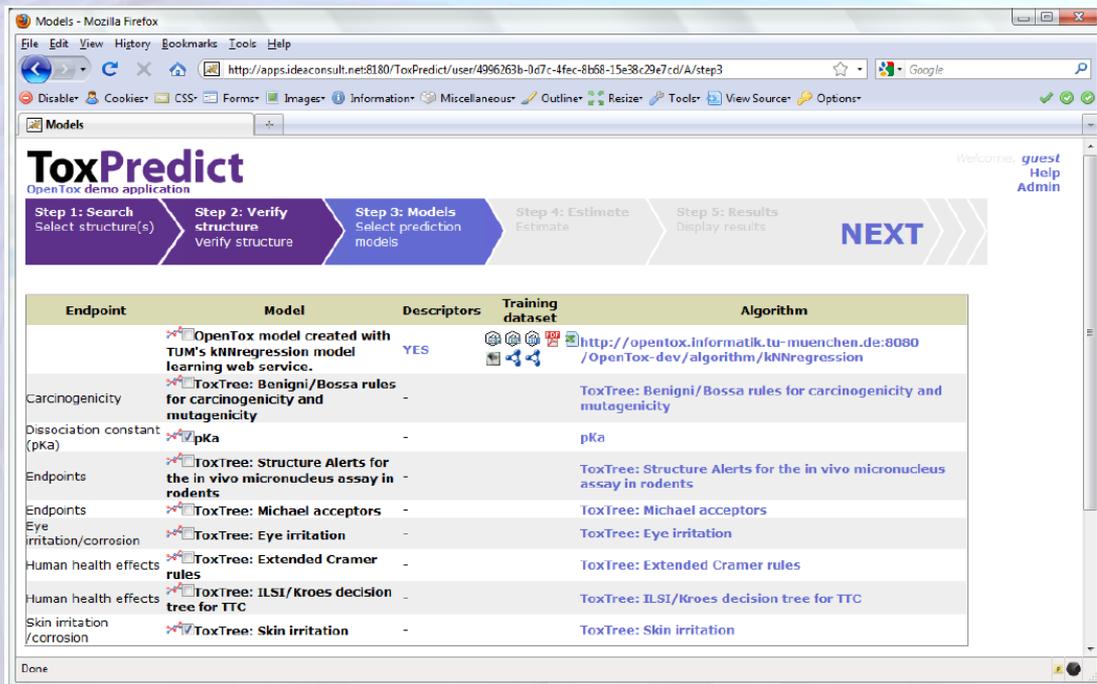
<http://myhost.com/model/predictivemodel1>

Prototype applications

- Two end user oriented demo applications, making use of OpenTox webservices, have been developed, deployed and are available for testing - toxcreate.org and toxpredict.org
- ToxCreat creates models from user supplied datasets
- ToxPredict uses existing OpenTox models to estimate chemical compound properties



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

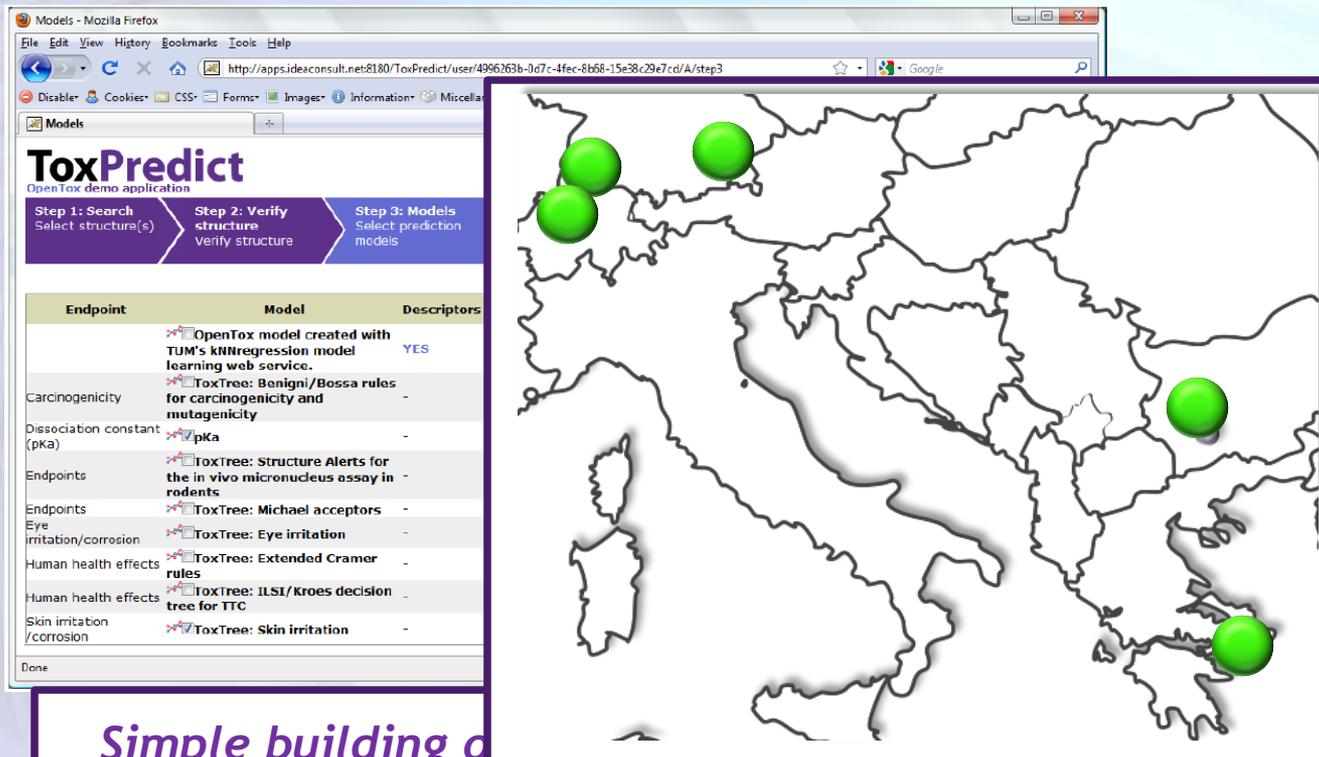
NEXT

Endpoint	Model	Descriptors	Training dataset	Algorithm
	OpenTox model created with TUM's kNNregression model learning web service.	YES	http://opentox.informatik.tu-muenchen.de:8080/OpenTox-dev/algorithm/kNNregression	
Carcinogenicity	ToxTree: Benigni/Bossa rules for carcinogenicity and mutagenicity	-		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

Done

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. The browser address bar displays the URL: <http://apps.ideaconsult.net:8180/ToxPredict/user/496263b-0d7c-4fec-8b68-15e38c29e7cd/A/step3>. The application title is "ToxPredict" and it is described as an "OpenTox demo application".

The interface features 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 this is a table with columns for "Endpoint", "Model", and "Descriptors".

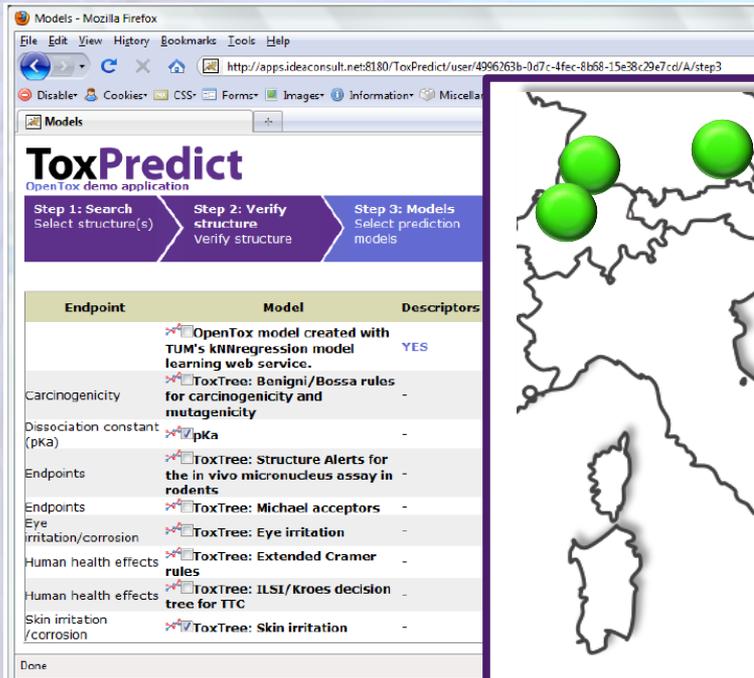
Endpoint	Model	Descriptors
	<input checked="" type="checkbox"/> OpenTox model created with TUM's kNNregression model learning web service.	YES
Carcinogenicity	<input checked="" type="checkbox"/> ToxTree: Benigni/Bossa rules for carcinogenicity and mutagenicity	-
Dissociation constant (pKa)	<input checked="" 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 status "Done" is visible. To the right of the browser window is a map of Europe with five green circular markers placed over various regions.

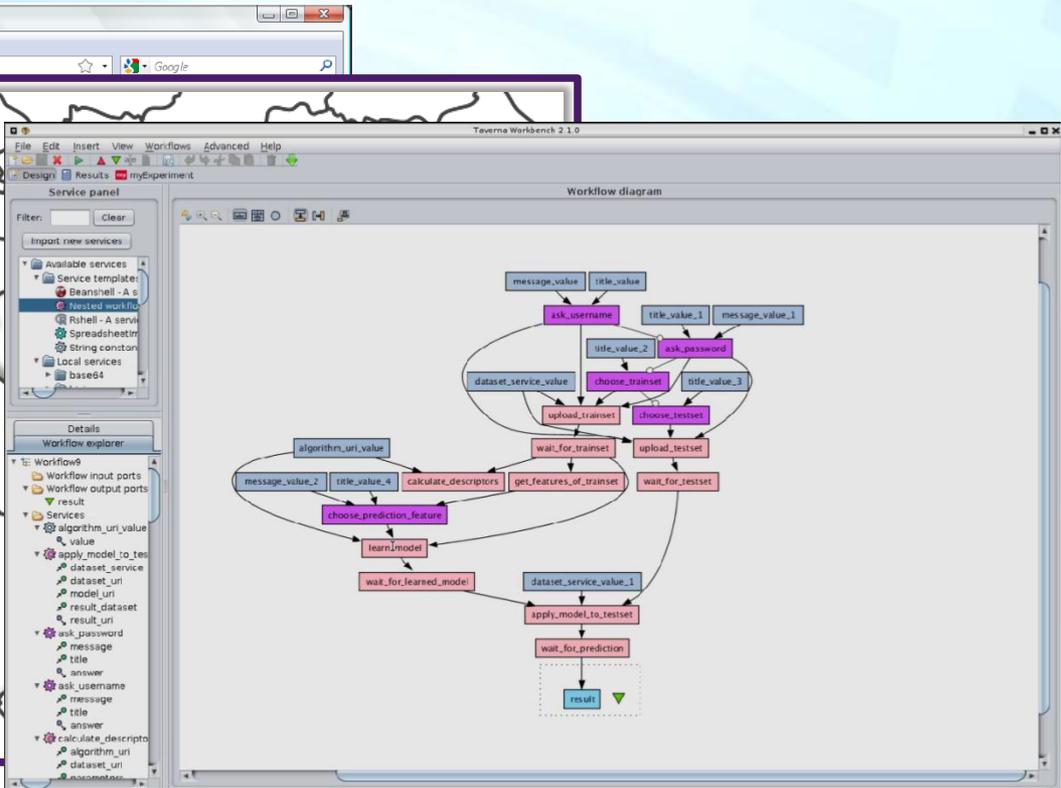
Simple building of applications methods and

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 Taverna Workbench interface displays a workflow diagram with the following steps:

- Input new services
- ask_username (message_value, title_value)
- ask_password (title_value_1, message_value_1)
- choose_trasnet (title_value_2)
- choose_testset (title_value_3)
- upload_trainset (dataset_service_value)
- upload_testset (choose_testset)
- wait_for_trainset
- wait_for_testset
- get_features_of_trainset
- calculate_descriptors (algorithm_uri_value)
- choose_prediction_feature (message_value_2, title_value_4)
- learn_model
- wait_for_learned_model
- dataset_service_value_1
- apply_model_to_testset
- wait_for_prediction
- result

Simple building of applications methods and

Distributed of wide range of methods

Integration into workflow systems for computational biology

OpenTox: Databases

Chemical compounds - Mozilla Firefox

File Edit View History Bookmarks Tools Help

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

Chemical compounds

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

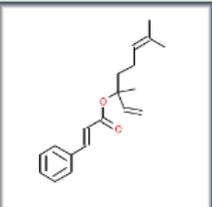
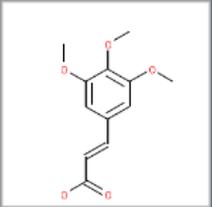
ambit

SMARTS

Keywords

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

Retrieve data SMARTS Download as Max number of hits:

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

Default
Identifiers
Datasets
Models
Endpoints
All descriptors
pKa
Molecule size
Electronic descriptors (PM3 optimized structure)
Electronic descriptors (original structure)
Toxtree: Cramer rules

http://apps.idea

Interoperability & Vocabulary

Dogs

- Collie
- Labrador



Cats

- Siamese
- Persian



Birds

- Sparrow
- Owl

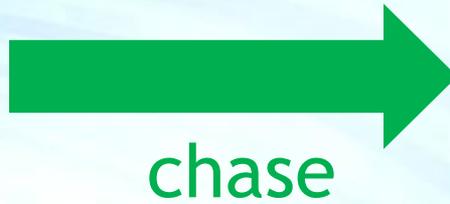


Interoperability & Vocabulary



Interoperability & Ontology

Org A

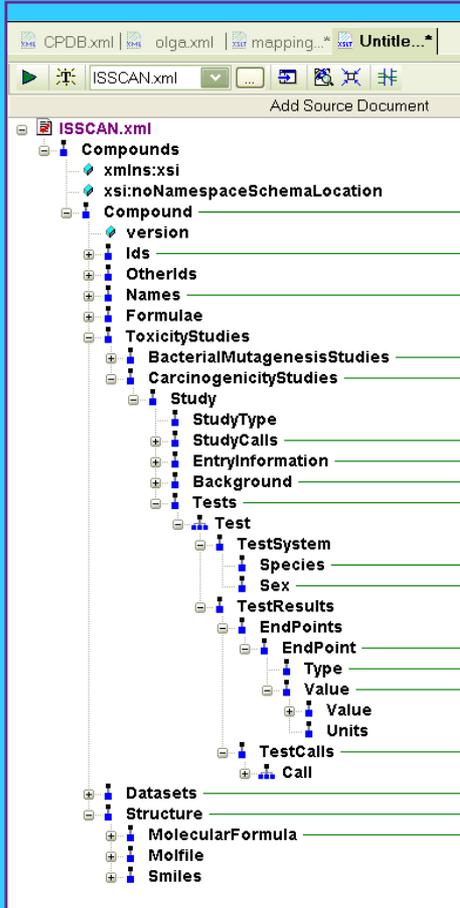


Org B



Toxicological Endpoint Ontology Development

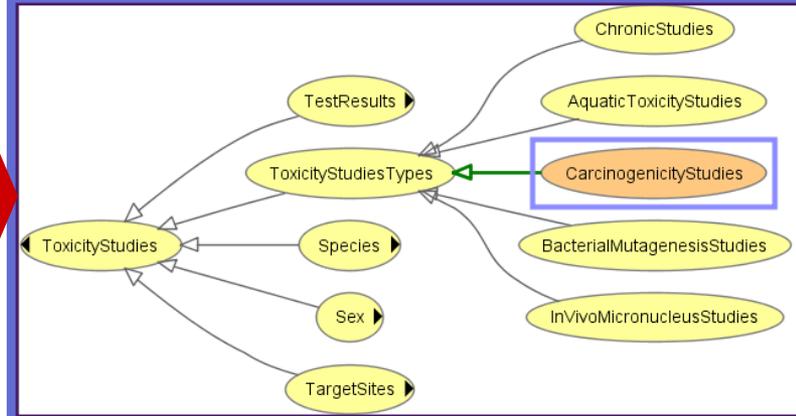
ToxML schema



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

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

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



Dataset and Ontology - find an assay, linked to specific gene

```
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 ?genename
where {
```

```
{?Feature dc:title ?title}.
```

```
{?Feature owl:sameAs ?assay}.
```

```
{?assay toxcast:gene ?geneid}.
```

```
{?assay toxcast:hasProperty ?genename}.
```

```
{?genename rdf:type toxcast:GENE_NAME}.
```

```
}
```

[http://ambit.uni-plovdiv.bg:8080/ambit2/dataset/961?feature_uris\[\]=http://ambit.uni-plovdiv.bg:8080/ambit2/feature/335126](http://ambit.uni-plovdiv.bg:8080/ambit2/dataset/961?feature_uris[]=http://ambit.uni-plovdiv.bg:8080/ambit2/feature/335126)



Feature	title	assay	geneid	genename
http://ambit.uni-plovdiv.bg:8080/ambit2/feature/335126	'ATO_RORE_CIS'	http://www.opentox.org/toxcast#ATO_RORE_CIS	http://bic2rdf.org/geneid/6095	http://www.opentox.org/toxcast#R0RA
http://ambit.uni-plovdiv.bg:8080/ambit2/feature/335187	'ATO_LXRb_TRANS'	http://www.opentox.org/toxcast#ATO_LXRb_TRANS	http://bic2rdf.org/geneid/7376	http://www.opentox.org/toxcast#R1H2
http://ambit.uni-plovdiv.bg:8080/ambit2/feature/335132	'ATO_MRE_CIS'	http://www.opentox.org/toxcast#ATO_MRE_CIS	http://bic2rdf.org/geneid/22823	http://www.opentox.org/toxcast#R1H7
http://ambit.uni-plovdiv.bg:8080/ambit2/feature/335060	'ATO_RARa_TRANS'	http://www.opentox.org/toxcast#ATO_RARa_TRANS	http://bic2rdf.org/geneid/5914	http://www.opentox.org/toxcast#R1RA
http://ambit.uni-plovdiv.bg:8080/ambit2/feature/335169	'ATO_Eta_CIS'	http://www.opentox.org/toxcast#ATO_Eta_CIS	http://bic2rdf.org/geneid/2113	http://www.opentox.org/toxcast#R1S1
http://ambit.uni-plovdiv.bg:8080/ambit2/feature/335090	'ATO_NF_kB_CIS'	http://www.opentox.org/toxcast#ATO_NF_kB_CIS	http://bic2rdf.org/geneid/4790	http://www.opentox.org/toxcast#R1KB1
http://ambit.uni-plovdiv.bg:8080/ambit2/feature/335136	'ATO_PPREM_CIS'	http://www.opentox.org/toxcast#ATO_PPREM_CIS	http://bic2rdf.org/geneid/9970	http://www.opentox.org/toxcast#R1R10

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

Dataset and Ontology - find an assay, linked to specific gene

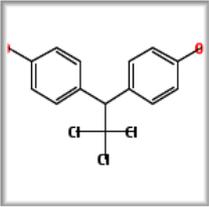
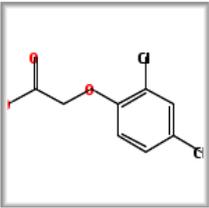
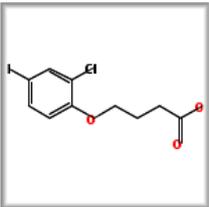
```

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}.
  {?assay toxcast:hasProperty ?genename}.
  {?genename rdf:type toxcast:GENE_NAME}.
}
    
```



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

#	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

?feat

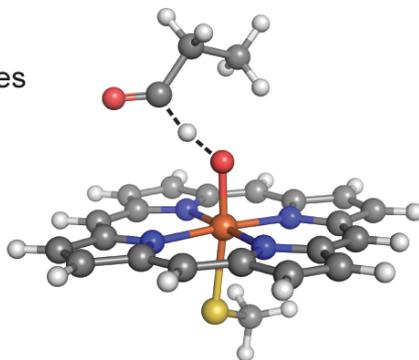
126

SMARTCyp Service for Predicting Metabolites

Atom Reactivity Library

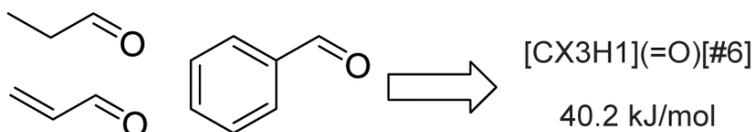
A. Calculate Quantum Chemical Reference Energies

Calculate transition state energies using density functional theory



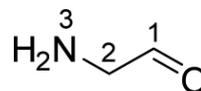
B. Define SMARTS Rules

Group calculations by fragments and calculate average energies



SMARTCyp

1. Assign Energies By SMARTS matching



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

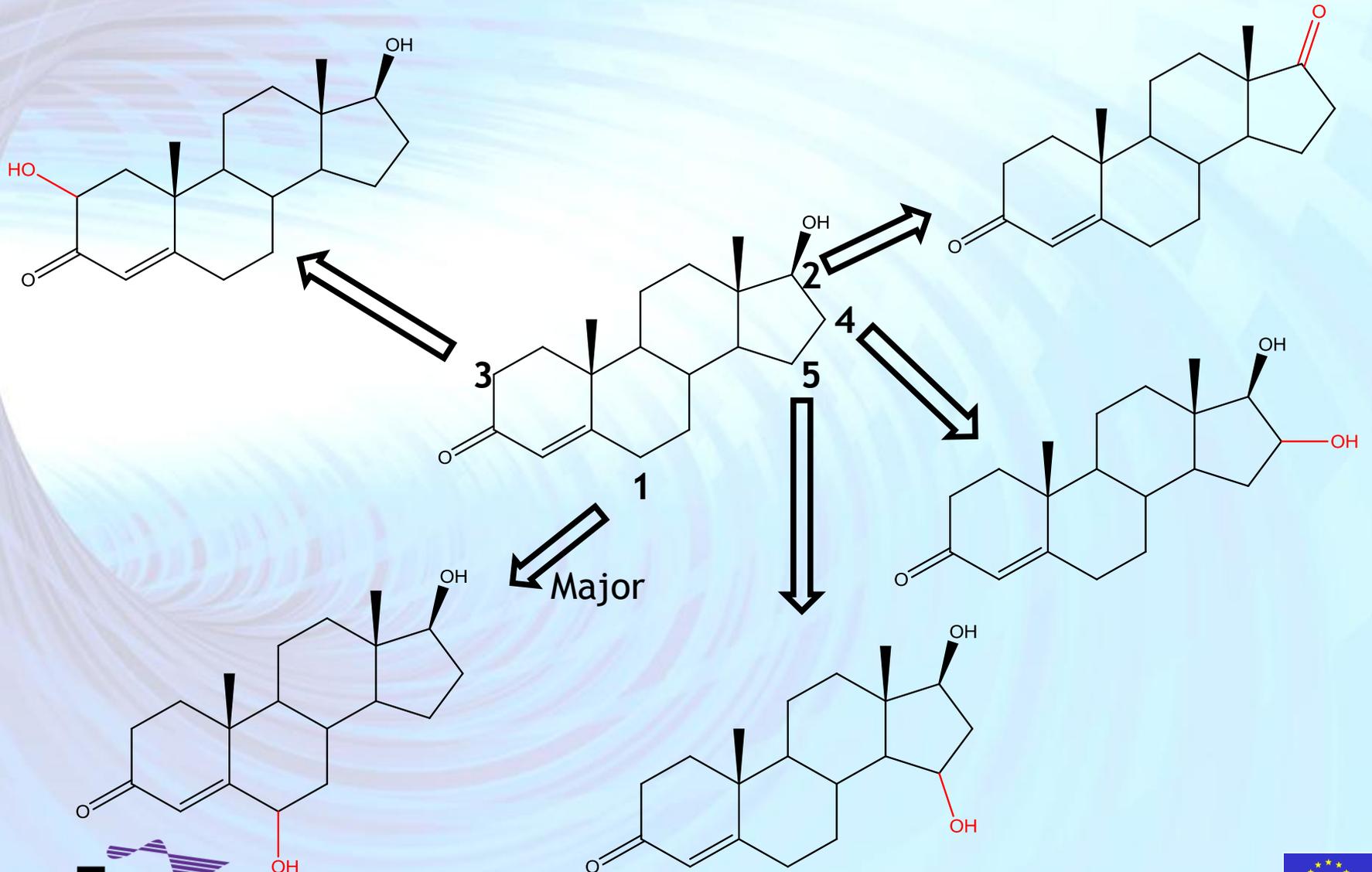
Atom 3 - Rank 3



SMARTCyp - developed by Patrik Rydberg, University of Copenhagen

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

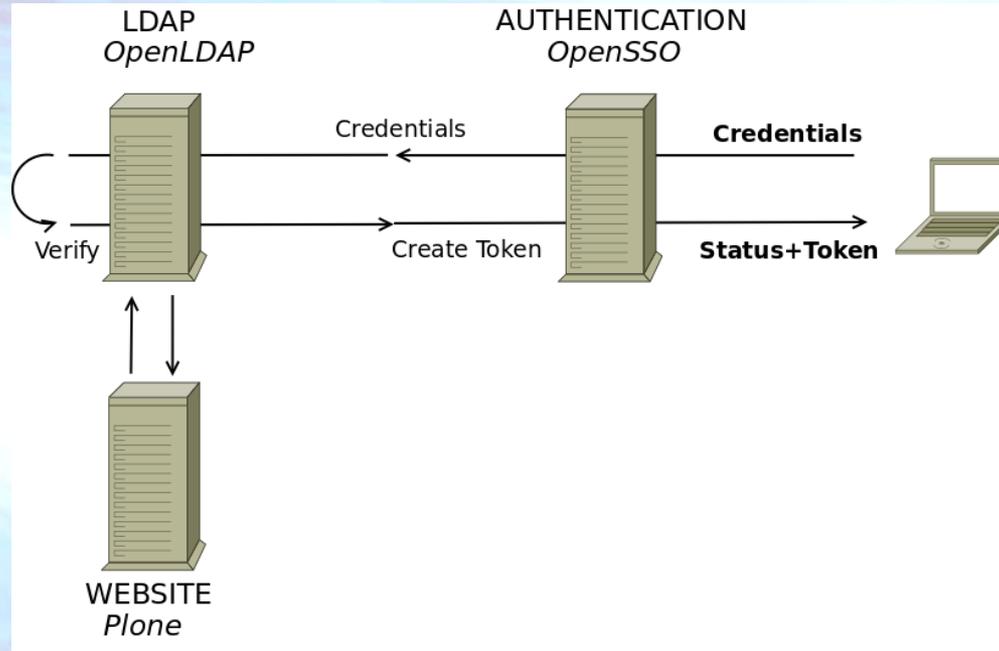
SmartCYP Prediction of Testosterone Metabolites



Controlling Access to Confidential Information

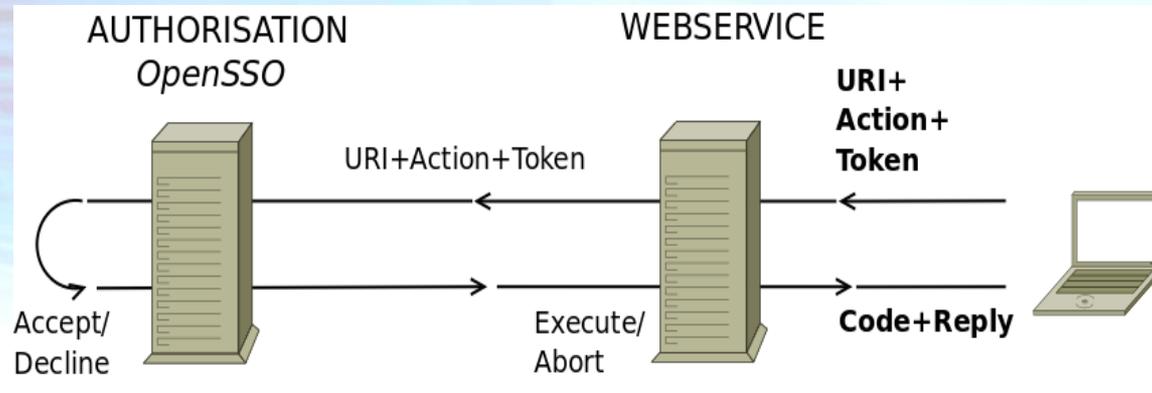
- OpenTox makes resources available through **URIs**
- OpenTox provides facilities to protect confidential information located at **URIs**. Two tasks are involved here:
 - **Authentication**: Confirming the identity of the user requesting access
 - **Authorisation**: Granting the confirmed identity access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



© Scott Adams, Inc./Dist. by UFS, Inc.

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', 'NMRShiftDB', 'OWL', 'PDF', 'SampleData', 'SampleData', 'Solubility', 'STATLite', 'Test', and 'Test2'. The main area contains a table with two rows of molecule data. The first row is selected and highlighted in blue. The Properties window on the right shows the 'General' tab with the following values:

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

OpenTox - ToxCast

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

Data Collection: EPA CCL3

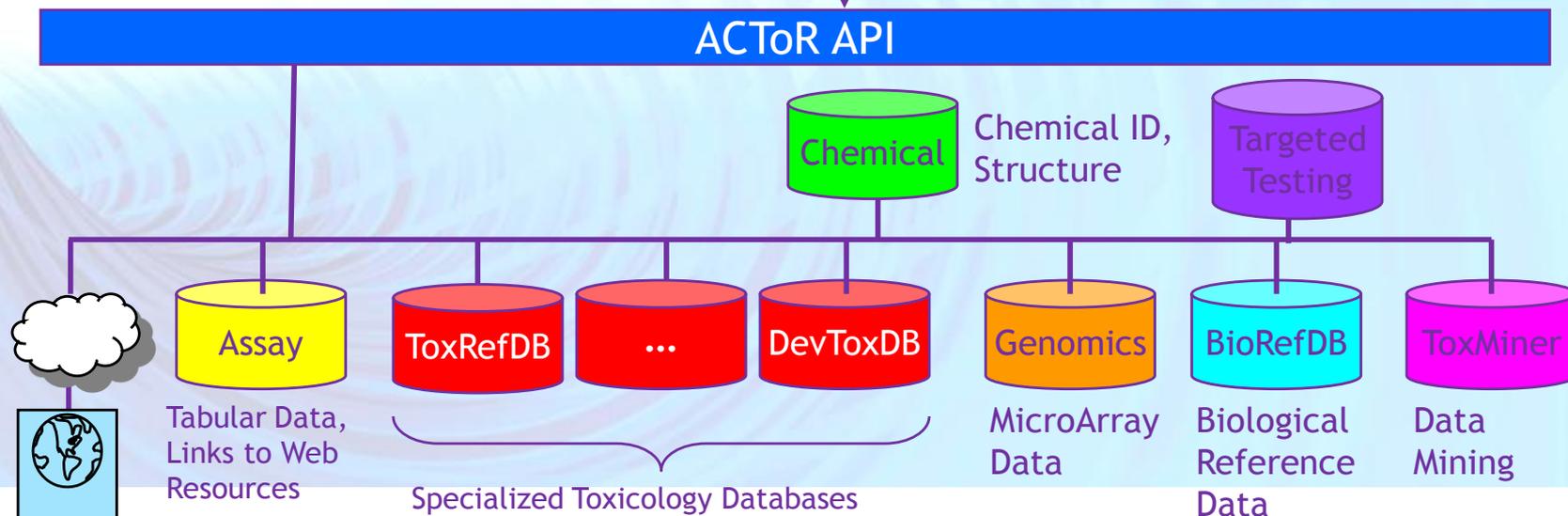
Name: EPA CCL3 List_03
Description: EPA has drinking water regulations for more than 90 contaminants. The Safe Drinking Water Act (SDWA) includes a process that we must follow to identify and list unregulated contaminants which may require a national drinking water regulation in the future. EPA must 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](#)

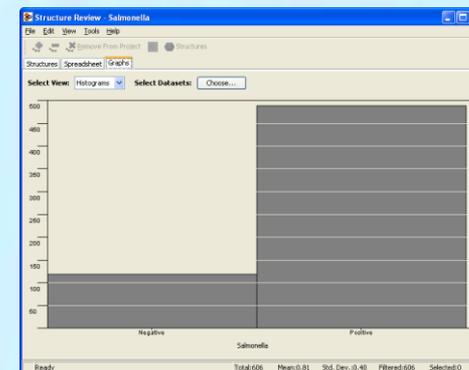
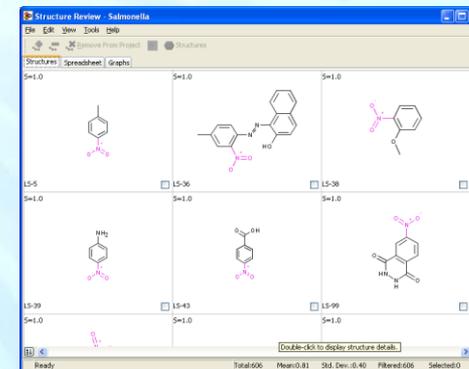
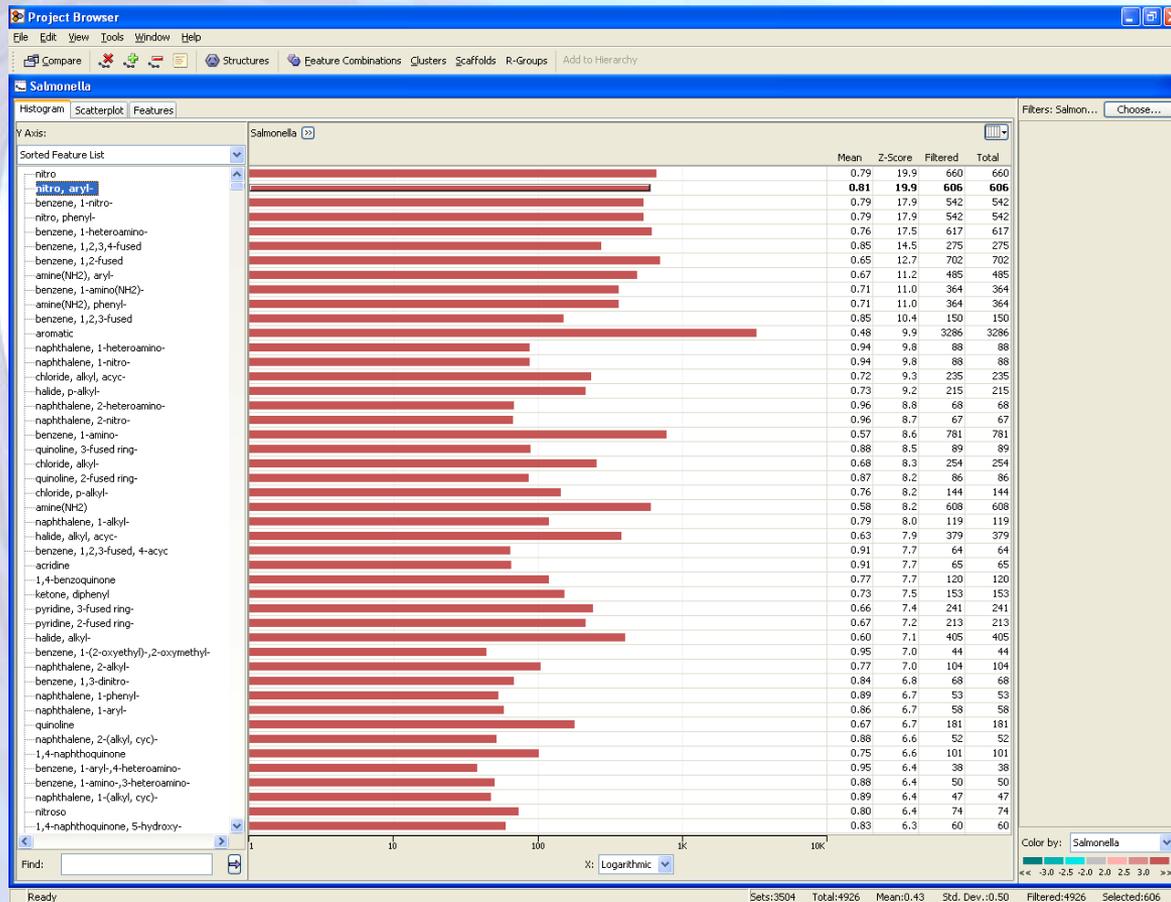
Structure	Name	CASRN	Generic Chemical Details	Toxicity					
				Hazard	Chemosensitivity	Genotoxicity	Developmental Toxicity	Reproductive Toxicity	Chemical Toxicity
<chem>ClC(Cl)(Cl)Cl</chem>	1,1,1,2-Tetrachloroethane	630-20-6	Details	Ha	Ca	G	D	R	Cr
<chem>CC(=O)Cl</chem>	1,1-Dichloroethane	75-34-3	Details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

ACToR Web Browser

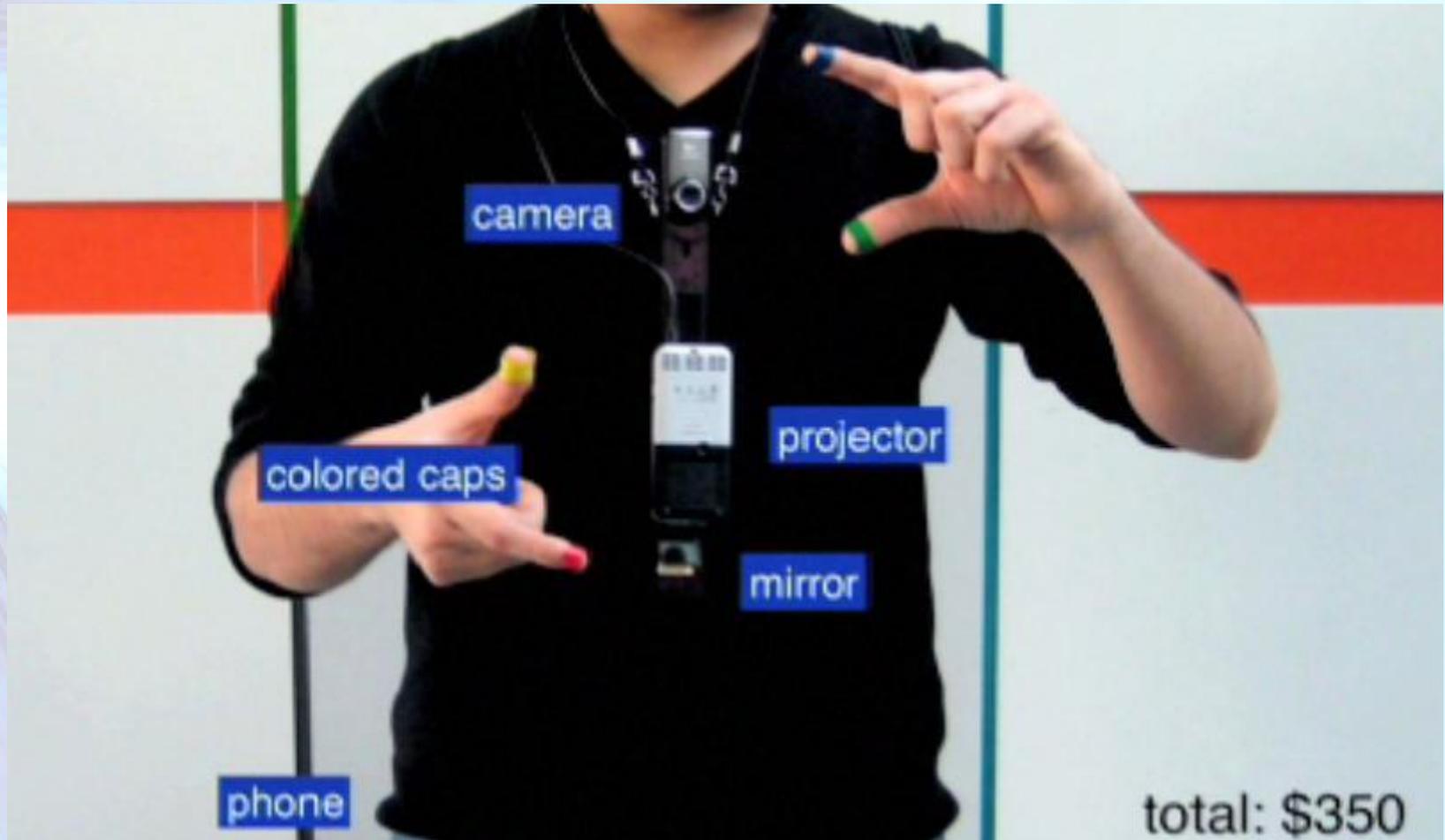


<http://actor.epa.gov/>

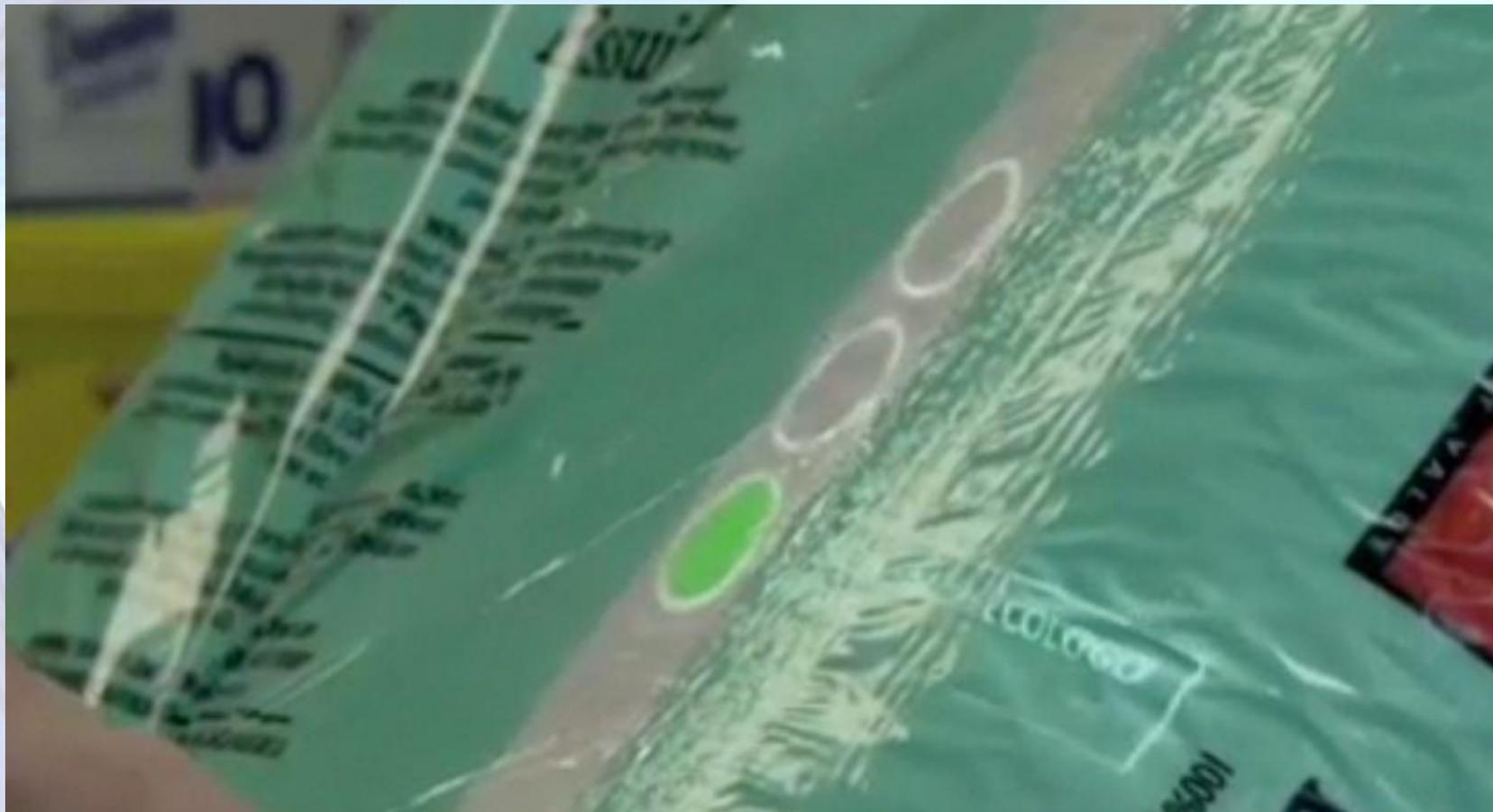
OpenTox - Leadscope



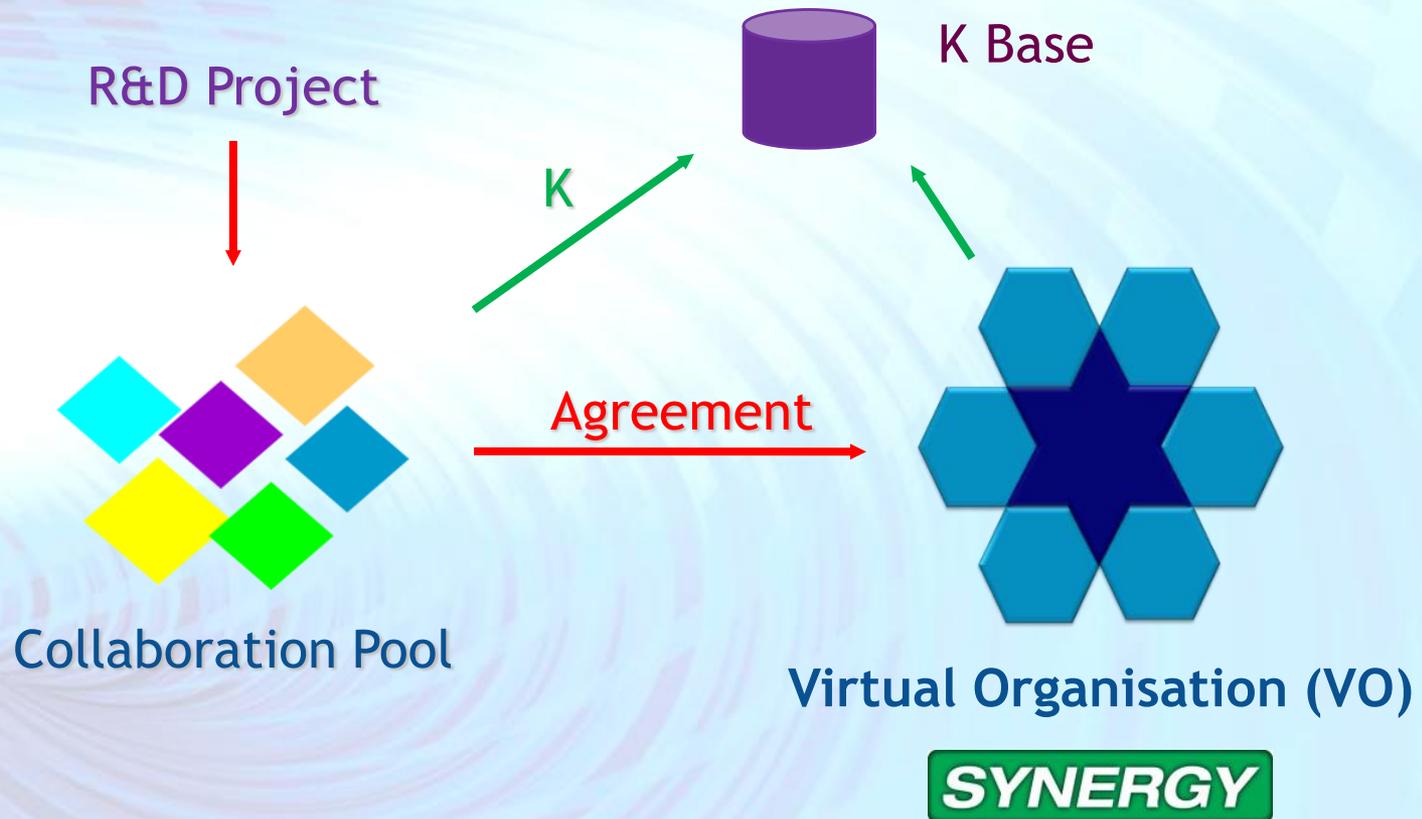
Augmented Reality



Processing Packaging Information



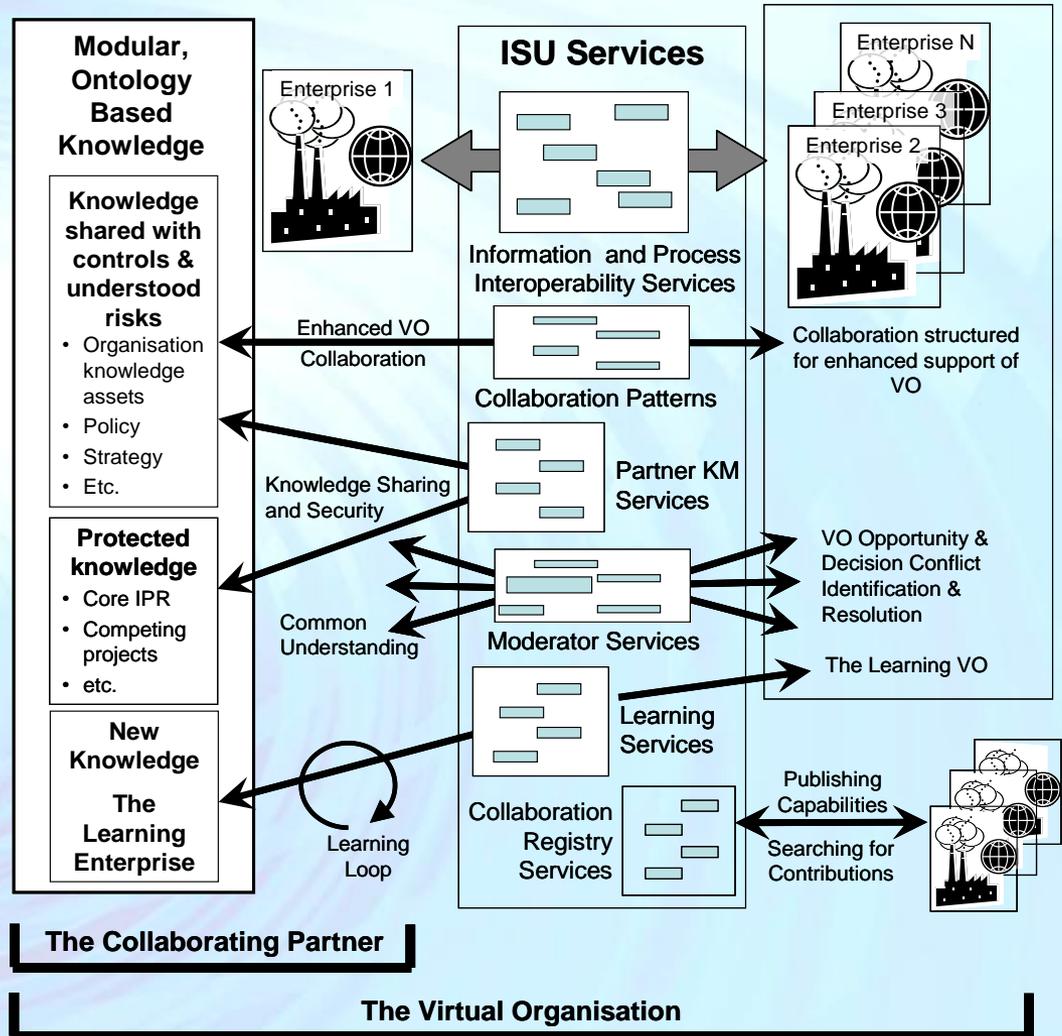
Virtual Organisation Pilots



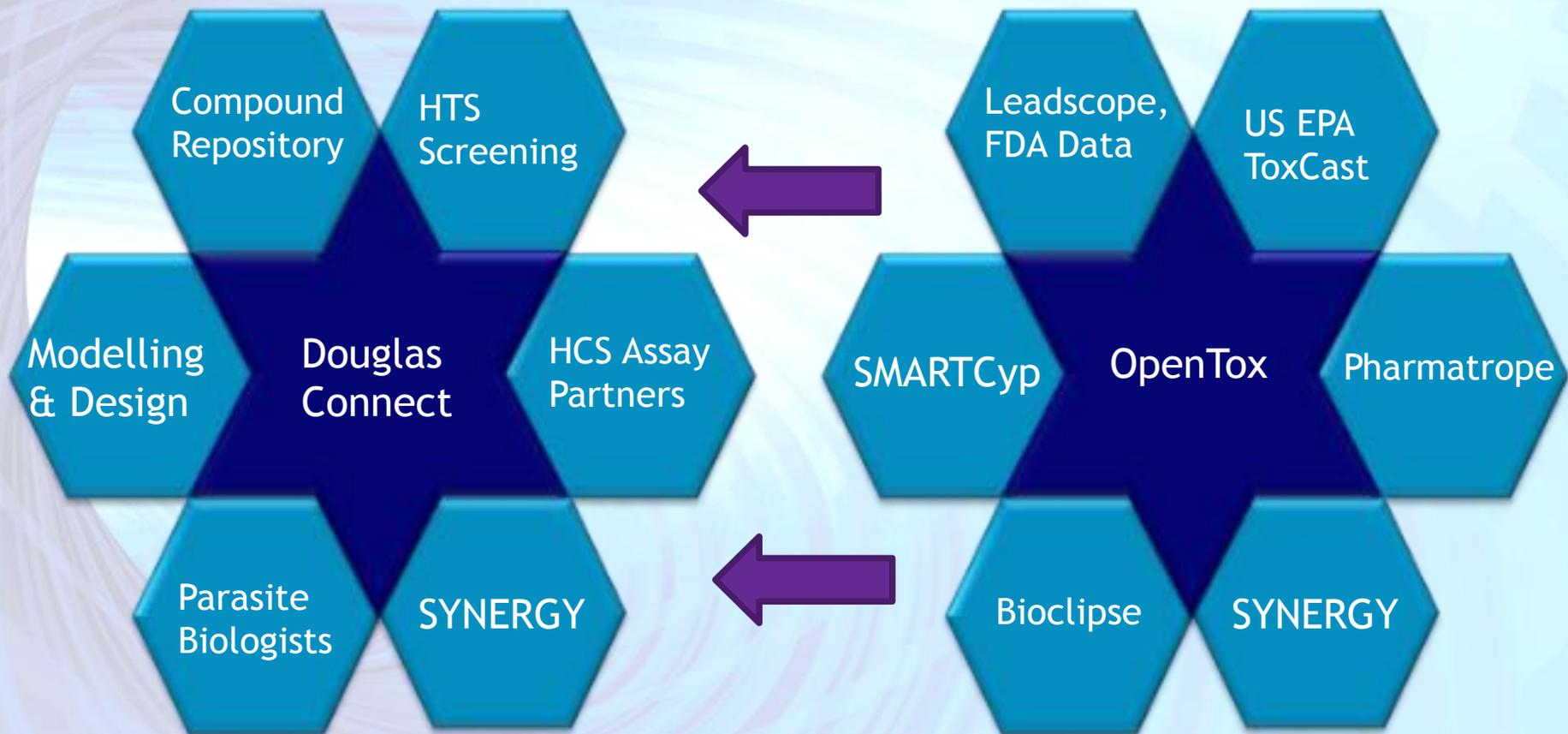
SYNERGY Collaboration Services for VOs



SYNERGY website:
www.synergy-ist.eu/



Virtual Organisation Pilots



Neglected Disease Drug Design VO

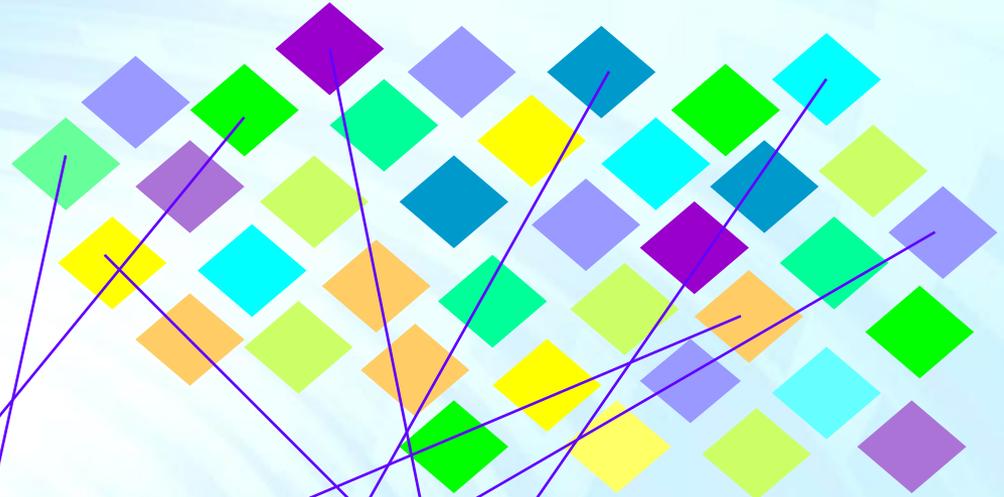
Predictive Toxicology VO

Creation of VO from Collaboration Pool

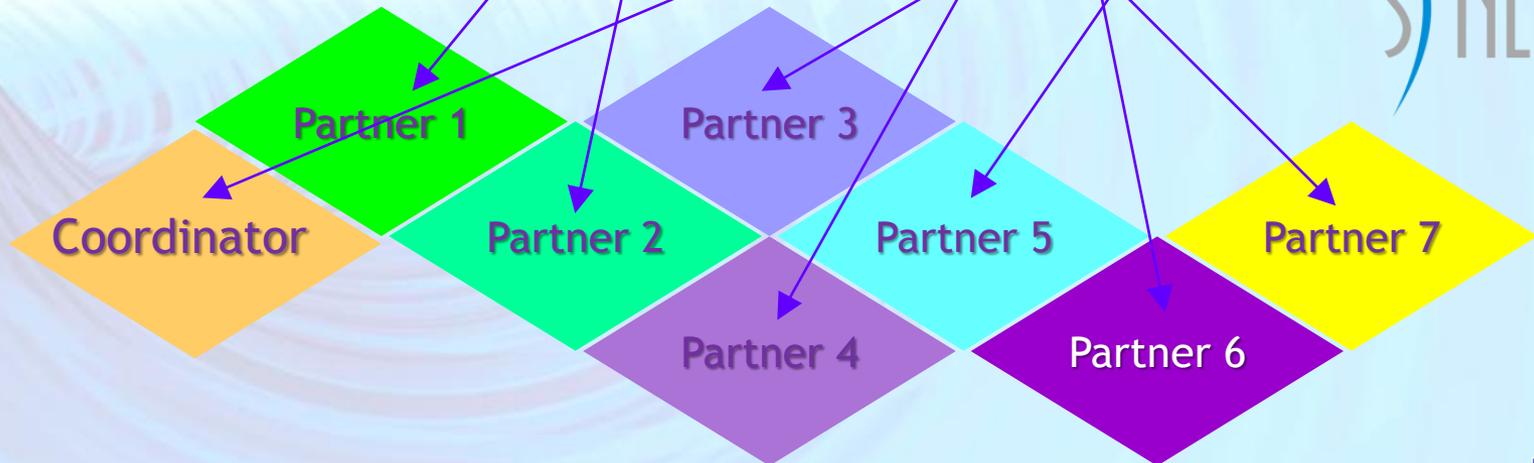
Network

Opportunity

Call for Tender
Need for joint effort
Major project



Virtual Organisation



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:

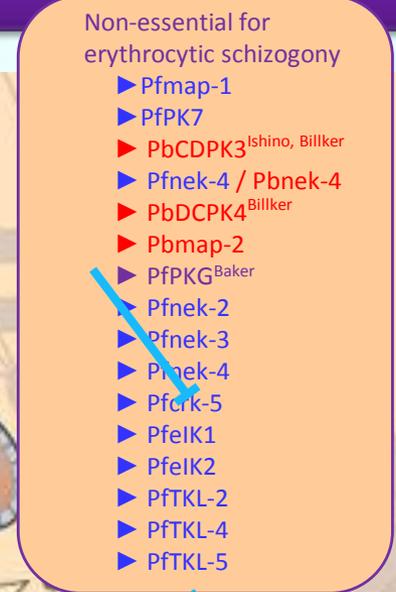
- ▶ PfPK7

Ookinete migration:

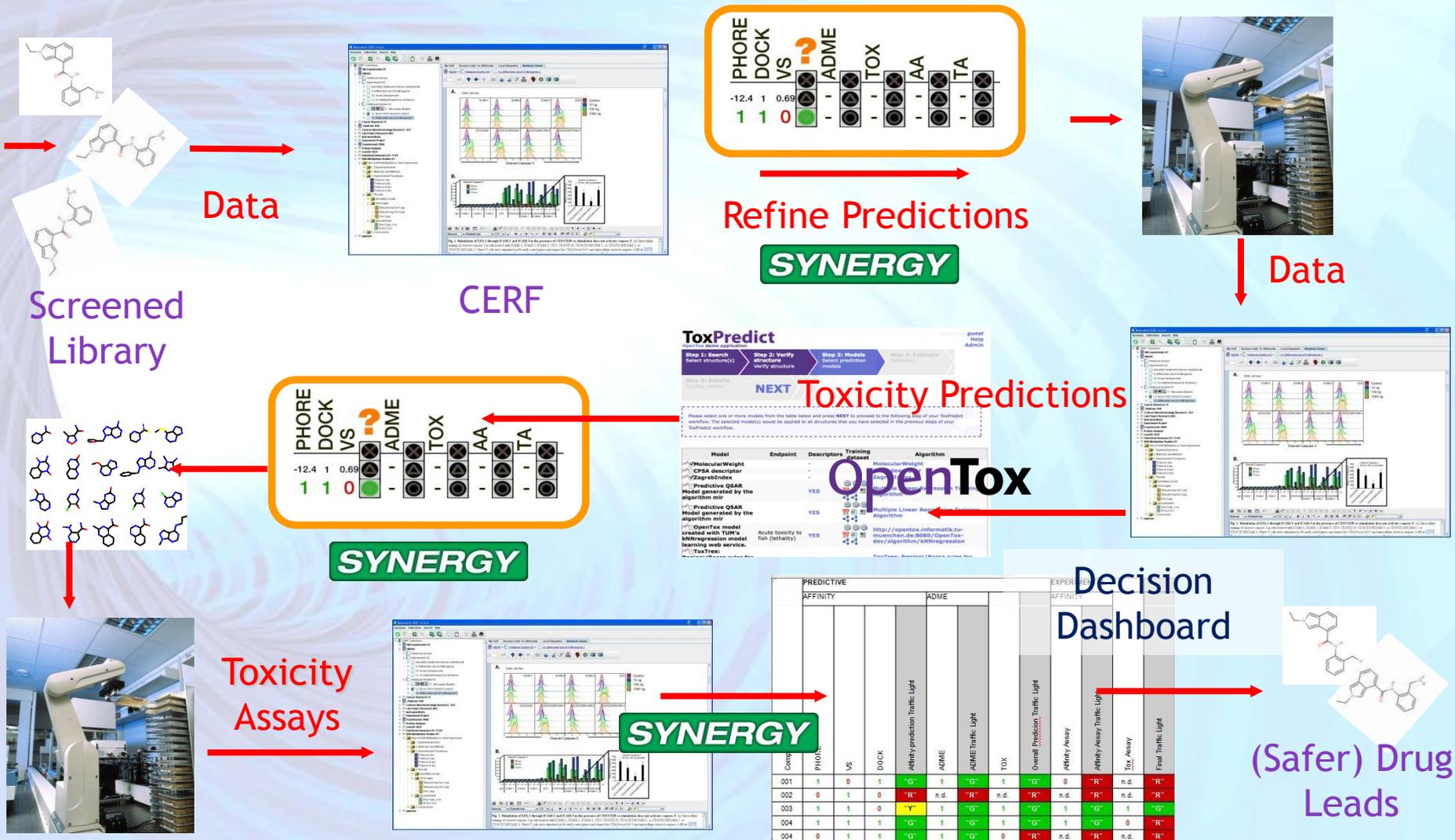
- ▶ **PbCDPK3^{Ishino, Billker}**

Zygote

Gametes



Synergy Drug Design Collaboration Pilot



Recording of Collaborative R&D

Rescentris CERF v2.6.0

Sessions Collections Search Help

My CERF Resource Info: 14. Differential Local Filesystem Notebook Viewer

NB260 > Notebook Section-03 > 14. Differential role of ICAM ligands-2

14

A. CD4+ (24 hrs)

B.

Control
10 ng
100 ng
1000 ng

Cleaved Caspase 3

Median Fluorescence

Cleaved Caspase 3
24 hrs
48 hrs
72 hrs

Cleaved Caspase-3
48 hrs, 100 ng stimulant

Normal Default font 12

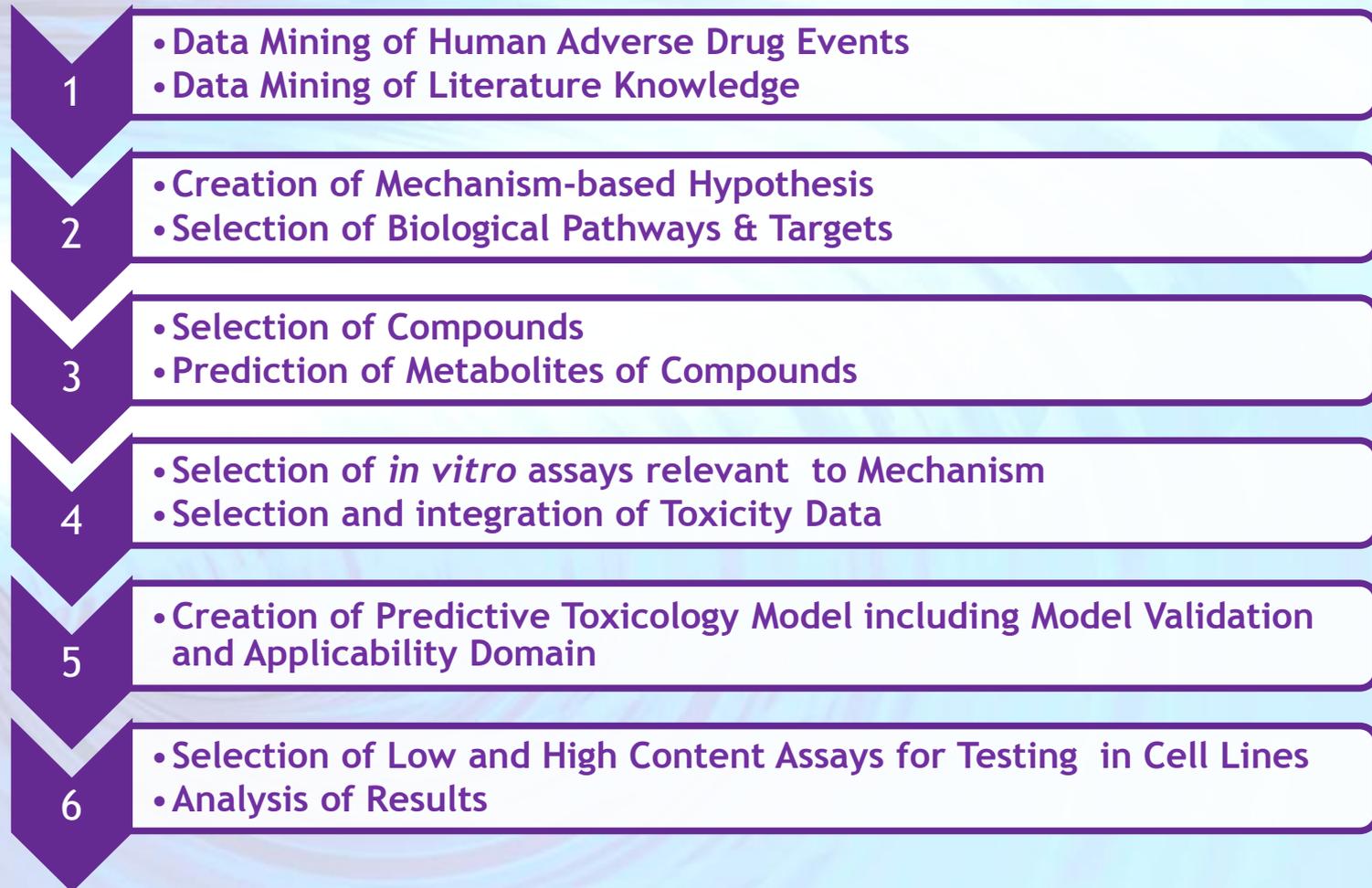
Fig. 1. Stimulation of LFA-1 through ICAM-2 and ICAM-3 in the presence of CD3/CD28 co-stimulation does not activate caspase-3. (A) Intracellular staining of cleaved-caspase 3 in cells treated with ICAM-1, ICAM-2, ICAM-3, CD3, CD3/CD28, CD3/CD28/ICAM-1, or CD3/CD28/ICAM-2, or CD3/CD28/ICAM-3. Naive T cells were stimulated in 96-well coated plates and stained for CD4-PerpCy5.5 and intracellular cleaved-caspase-3-FE at 24 hrs.

Controlled
Vocabularies

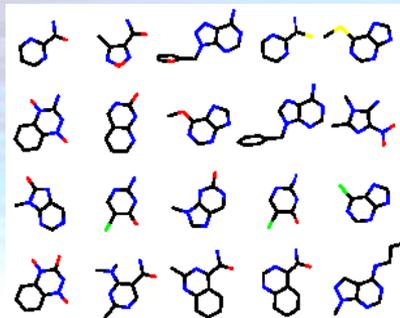
Visualisation

Collaborative
Electronic
Laboratory
Notebook (ELN)

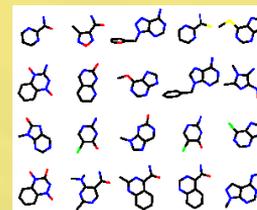
OpenTox - Synergy Predictive Toxicology VO Pilot Strategy Development & Case Study



1. A library of compounds is entered to the ELN



ELN

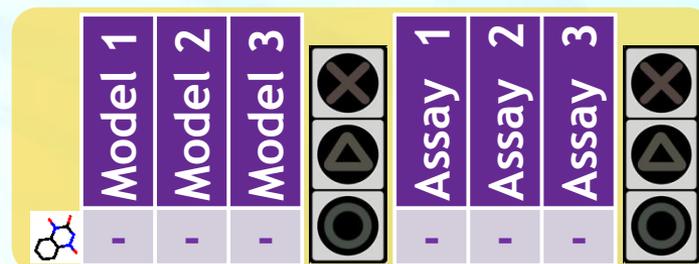


Synergy

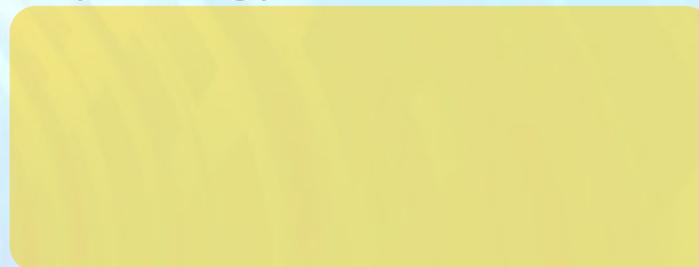
OpenTox

2. Each compound is assigned a data structure in ELN

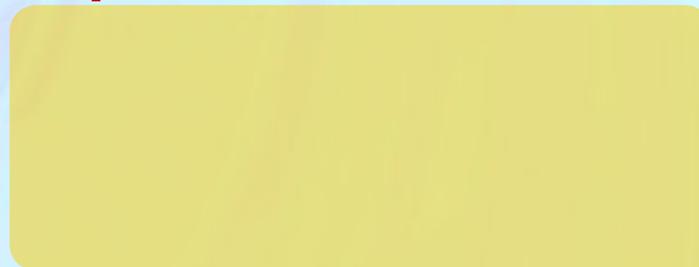
ELN



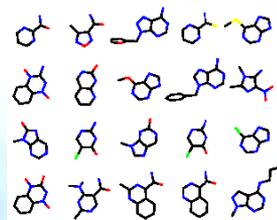
Synergy



OpenTox



3. ELN passes compounds to OpenTox and SYNERGY



ELN

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

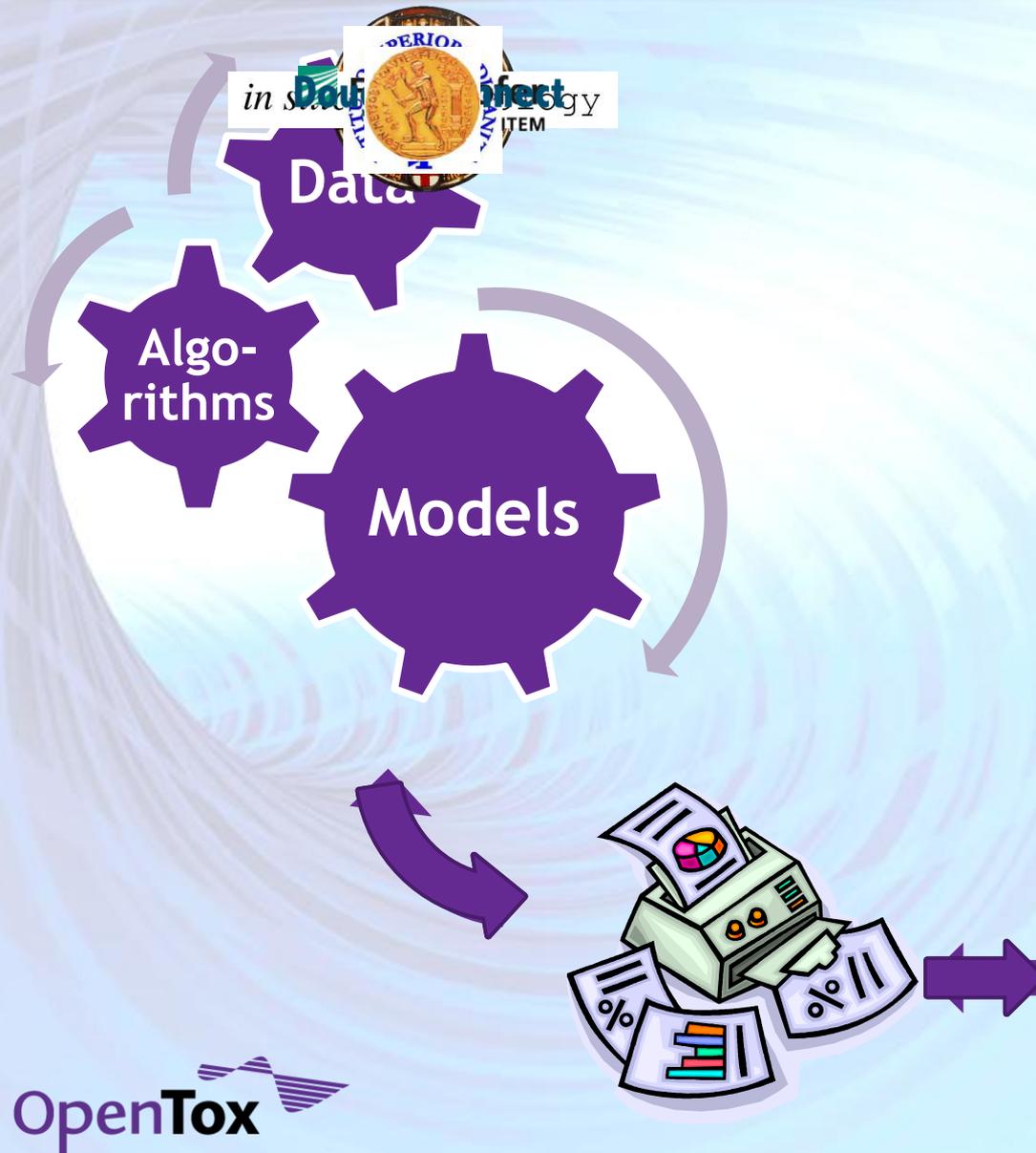
Synergy

	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		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	
	1	0	1		-	-	-	

5. OpenTox sends back a report to ELN



ELN

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

Synergy

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

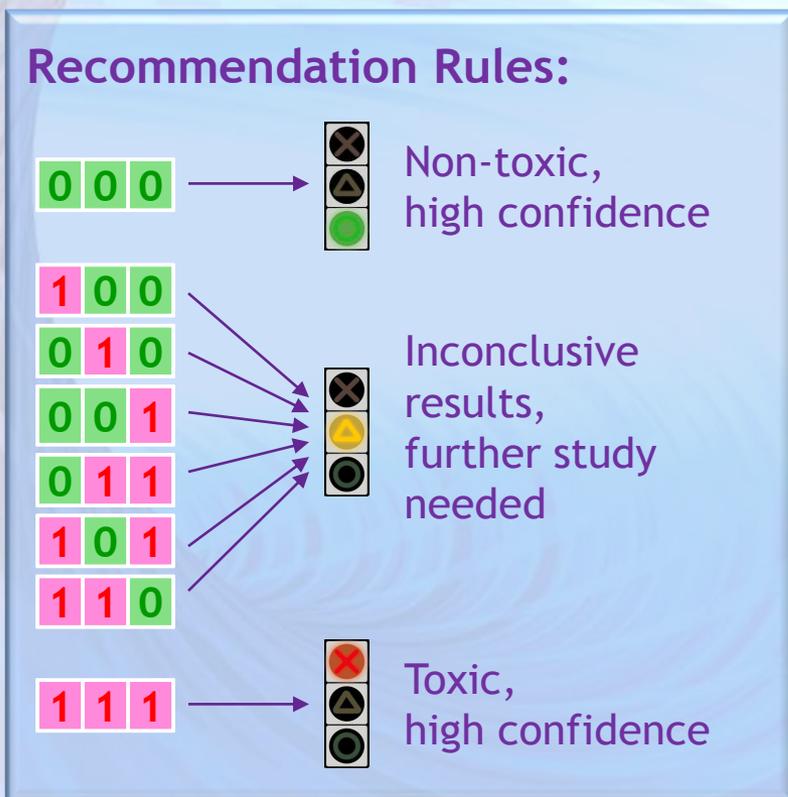
OpenTox

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

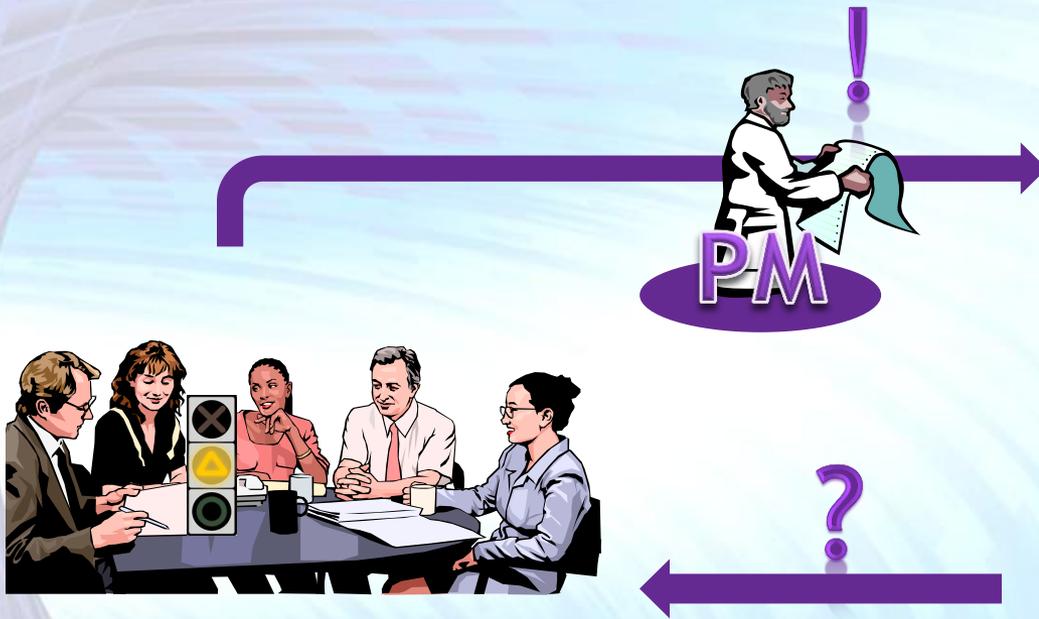
6. ELN sends the results to SYNERGY



7. SYNERGY applies the Recommendation Rules



8. Inconclusive data → SYNERGY calls a meeting



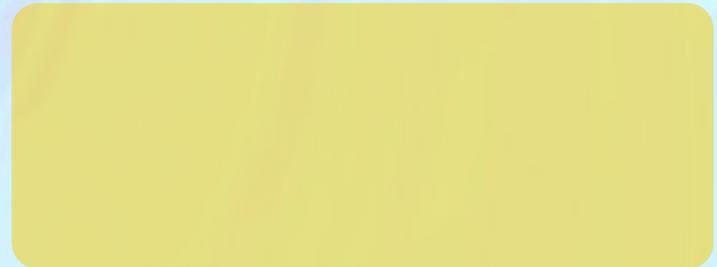
ELN

	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



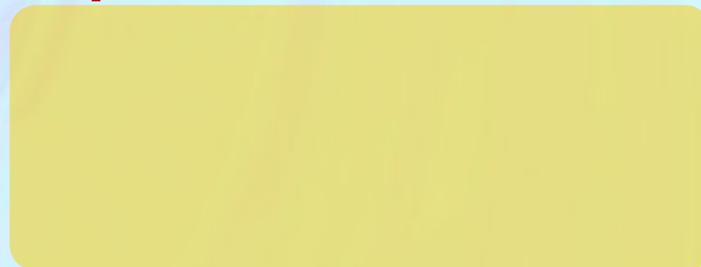
ELN

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

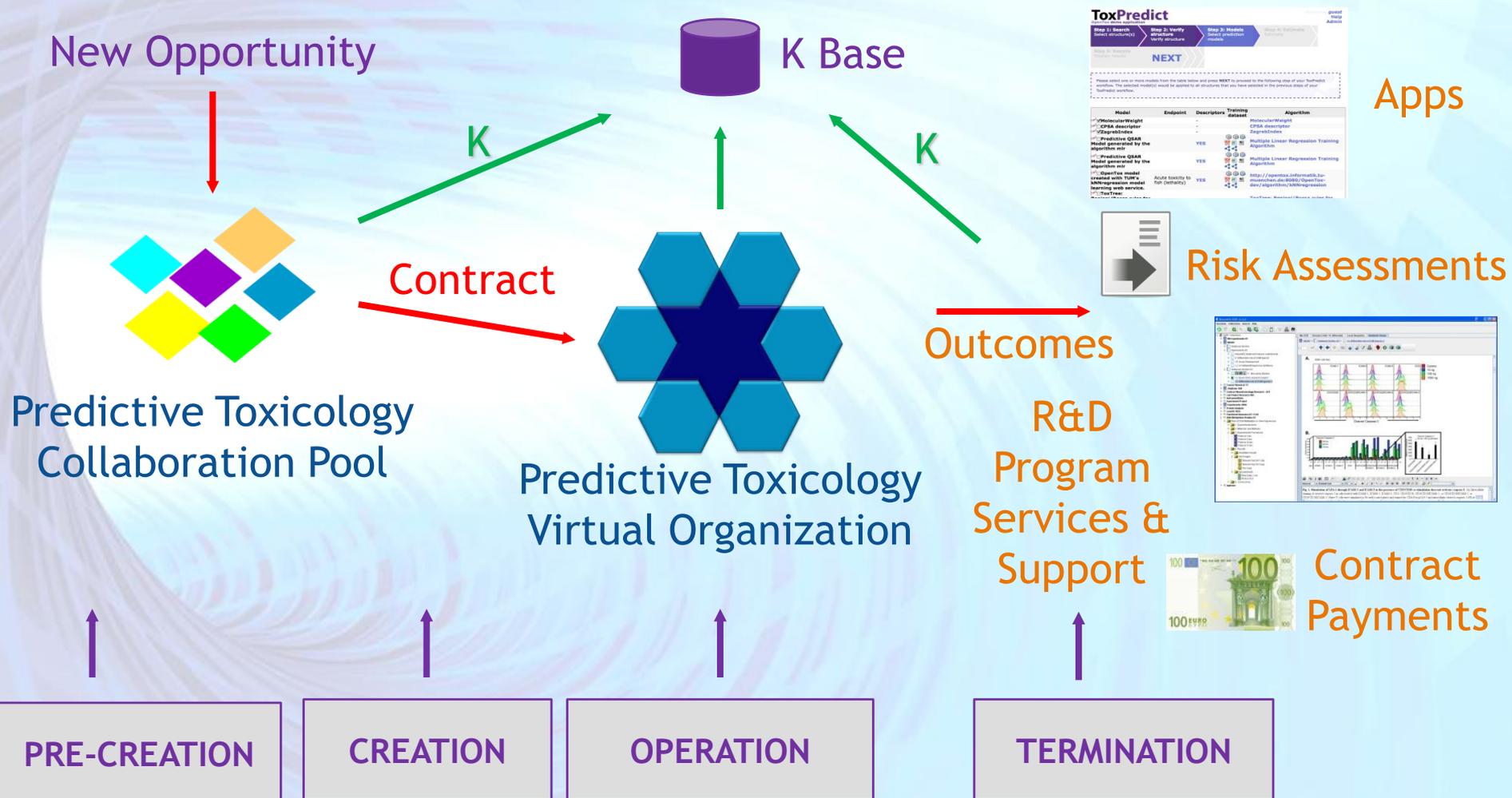
Synergy

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

OpenTox



Sustainability Model



Service Support of Virtual Organisation Lifecycle

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



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

Final words...

For more information, visit

www.opentox.org

Contact me:

barry.hardy@douglasconnect.com

**Many thanks for your
attention!**



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).