Deliverable D1.3

Prototype evaluation, improved API and interface designs

Grant Agreement	Health-F5-2008-200787
Acronym	OpenTox
Name	An Open Source Predictive Toxicology Framework
Coordinator	Douglas Connect

Contract No.	Health-F5-2008-200787	
Document Type:	Deliverable Report	
WP/Task:	WP1 / D1.3	
Name	Prototype evaluation, improved API, and interface designs	
Document ID:	OpenTox Deliverable Report WP1-D1.3	
Date:	Sept 30, 2010	
Status:	Final Version	
Organisation:	IST	
Contributors	Andreas Maunz Vedrin Jeliazkov Sylvia Escher Barry Hardy	IST IDEA ITEM DC

Distribution:	Public

Purpose of Document:	To document and disseminate results for this deliverable

Document History:	1 – First draft circulated on Aug 30, 2010 2 – First draft reviewed and edited by VJ on Aug 31, 2010 3 – Edited by SE on Sep 03, 2010 4 – Final version, Sept 30, 2010

Table of Contents

Summary

Two initial OpenTox prototype applications were evaluated: ToxPredict (toxpredict.org) which predicts and reports on toxicities for endpoints for an input chemical structure, and ToxCreate (toxcreate.org) which builds and validates a predictive toxicity model based on an input toxicology dataset. Templates were created for the recording of user feedback and beta testing results; initial testing provided useful guidance for subsequent development which is ongoing and will be subject to further testing as OpenTox evolves towards its final prototype applications in 2011.

The OpenTox Application Programming Interface (API) published openly at opentox.org/dev/apis has already found interest in the cheminformatics and bioinformatics communities. Currently, integrations into several different software packages are under development. We have received feedback and demand for new features from these collaborating developer communities.

A platform for continuous availability and performance monitoring of selected OpenTox web services has been designed and deployed. We discuss the rationale for such monitoring and provide an overview of the results obtained so far.

Based on initial evaluation and discussion of the OpenTox API in 2009, the API 1.1 was developed and released in late 2009, and included a commitment to the semantic representation of all OpenTox resources. Experiences with the API 1.1 during prototyping has led to numerous modifications and improvements which have been included in version 1.2 due for release in late 2010.

Acknowledgements

Research Funding

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).

Project Partners

Douglas Connect (DC), In Silico Toxicology (IST), Ideaconsult (IDEA), Istituto Superiore di Sanita' (ISS), Technical University of Munich (TUM), Albert Ludwigs University Freiburg (ALU), National Technical University of Athens (NTUA), David Gallagher (DG), Institute of Biomedical Chemistry of the Russian Academy of Medical Sciences (IBMC), Seascape Learning (SL), Jawaharlal Nehru University (JNU), Fraunhofer Institute for Toxicology & Experimental Medicine (ITEM).

Advisory Board

European Centre for the Validation of Alternative Methods, European Joint Research Centre, U.S Environmental Protection Agency, U.S. Food & Drug Administration, Nestlé, Roche, AstraZeneca, Lhasa, Leadscope, University of North Carolina, Pharmatrope, Bioclipse, EC Environment Directorate General, Organisation for Economic Co-operation & Development, CADASTER, Bayer Healthcare.

Correspondence

Dr. Barry Hardy, OpenTox Project Coordinator and Director, Community of Practice & Research Activities, Douglas Connect, Baermeggenweg 14, 4314 Zeiningen, Switzerland

Email: barry.hardy -(at)- douglasconnect.com

1. Introduction

Initial design work within OpenTox resulted in the conceptual and technical framework described and discussed by the OpenTox development community (opentox.org/dev) and which in turn led to the specifications of the OpenTox Application Programming Interfaces (APIs) published openly at opentox.org/dev/apis starting in 2009. This in turn led to the development of initial OpenTox web services and two initial prototype applications combining web services of different types into satisfying two core Use Cases: ToxPredict (toxpredict.org) which predicts and reports on toxicities for endpoints for an input chemical structure, and ToxCreate (toxcreate.org) which builds and validates a predictive toxicity model based on an input toxicology dataset.

In this work we carried out an evaluation of the initial OpenTox APIs, the prototype applications and initial feedback from other development groups who have commenced work to interoperate their software with OpenTox. The impact of the evaluation results on OpenTox specifications was discussed and specifications and development plans modified to take advantage of lessons learned from initial experiences.

The API (1.2) was extended by services to secure confidential data within OpenTox. Specifically, authentication and authorization interfaces were developed and implemented. We report the complete draft API in tabulated form.

2. Prototype Evaluation

2.1 Web Application Beta Testing

Beta testing templates were developed for ToxPredict and ToxCreate which are provided as sample completed forms in Appendices A and B. The beta testing results demonstrated that the prototypes completed the corresponding core use cases successfully but, as not unexpected, that many bugs and issues could be found for further resolution.

2.2 Continuous availability and performance monitoring

In this section we discuss the rationale for continuous availability and performance monitoring of OpenTox web services and provide a summary of the results obtained so far.

The OpenTox APIs have been designed, implemented and optimized regarding a range of important software quality metrics such as:

1. Performance – ability of web services to process tasks quickly;
2. Scalability – ability of web services to respond to a high number of concurrent requests and/or process tasks involving vast amounts of data without severe impact on response time;
3. Efficiency – fulfillment of purpose without waste of resources, such as memory, space and processor utilization, network bandwidth, time, etc.;
4. Reliability – ability to be expected to perform its intended functions satisfactorily;
5. Usability – convenience and practicality of use;
6. Maintainability – propensity to facilitate updates to satisfy new requirements;
7. Portability – ability to be run well and easily on multiple computer configurations;
8. Interoperability – ability to integrate seamlessly new API-compliant web services;

We selected SmokePing (oss.oetiker.ch/smokeping/) for monitoring in order to obtain automated measurements and statistics for the first three metrics (performance, scalability and efficiency). As a side effect of this continuous monitoring, we also have an early warning mechanism for web service outages. The monitoring results can be accessed online at ambit.uni-plovdiv.bg/cgi-bin/smokeping.cgi

We use a custom modified cURL (curl.haxx.se) probe for querying selected targets of OpenTox API-compliant web services. Figure 1 illustrates the graphical output of these tests. The chart on the Figure presents a summary of the test results for the last 24 hours, but of course, one could look at similar charts for arbitrary time periods when tests have been running. Some important observations could be made regarding the results in the Figure:

- the response time for most of the queries is lower than 20 ms – this demonstrates clearly that the tested web service performs well regarding the first software quality metric (ability to process tasks quickly);
- some of the queries exhibit much longer response time (e.g. StructDiagCSLS – 912 ms, StructDiagDaylight – 933 ms); StructSDFPubChem is the worst case with a response time of 6200 ms; what all these targets have in common is that they rely on 3rd party services, which are often overloaded and have several orders of magnitudes longer response time;
- the loss per target is zero for most of the targets; this means that all concurrent requests are processed successfully within the configured timeout;
- some of the queries (StructDiagCSLS, StructDiagDaylight2, StructSDFPubChem) exhibit some loss; again, all these are targets that rely on 3rd party services;
- a small number of targets (e.g. J48ModCreation, AllFPerCmpd, ApplypKaModel) could be optimized further in order to lower their response time and bring them closer to the ones exhibited by the vast

majority of targets; however, this might be not possible in some of the cases, since the processing involved might require more time to complete.

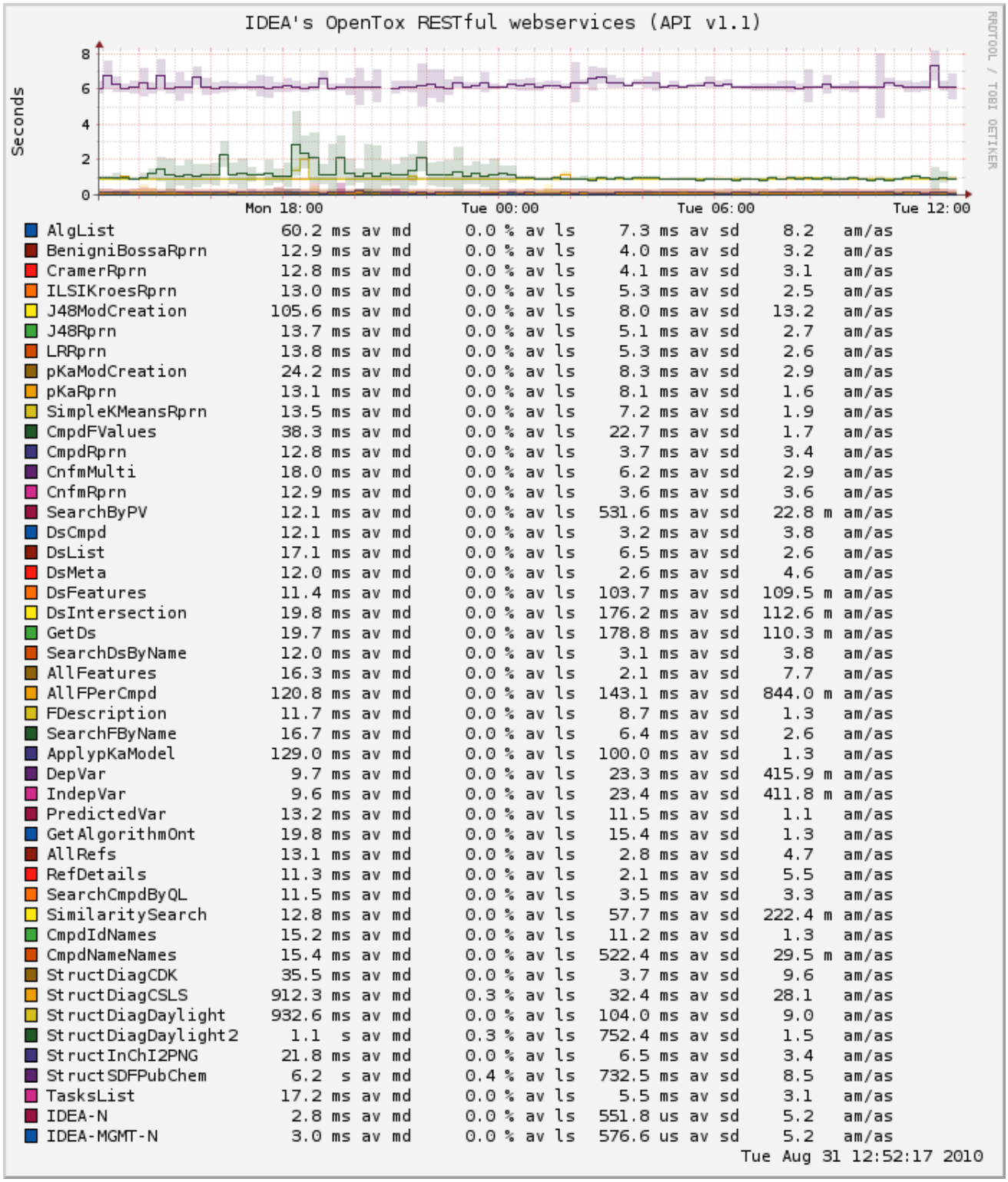


Figure 1 Visualisation of SmokePing monitoring results for IDEA OpenTox web services. The first column on the left lists the different targets being tested. The second column lists the average median latency per target. The third column lists the average loss per target (in this case loss occurs if no response is received back from the web service within a given timeout). The fourth column lists the average standard deviation of the multiple measurements in each round per target. And the last

column on the right represents the ratio of average median and average standard deviation per target. Tests are run every 5 minutes with 10 queries per target being sent.

This monitoring has helped us to identify and remove a number of performance bottlenecks, scalability and efficiency issues (e.g. memory leaks) in monitored services. As an example, consider the following figure, presenting detailed statistics for the TasksList target for the last 90 days.

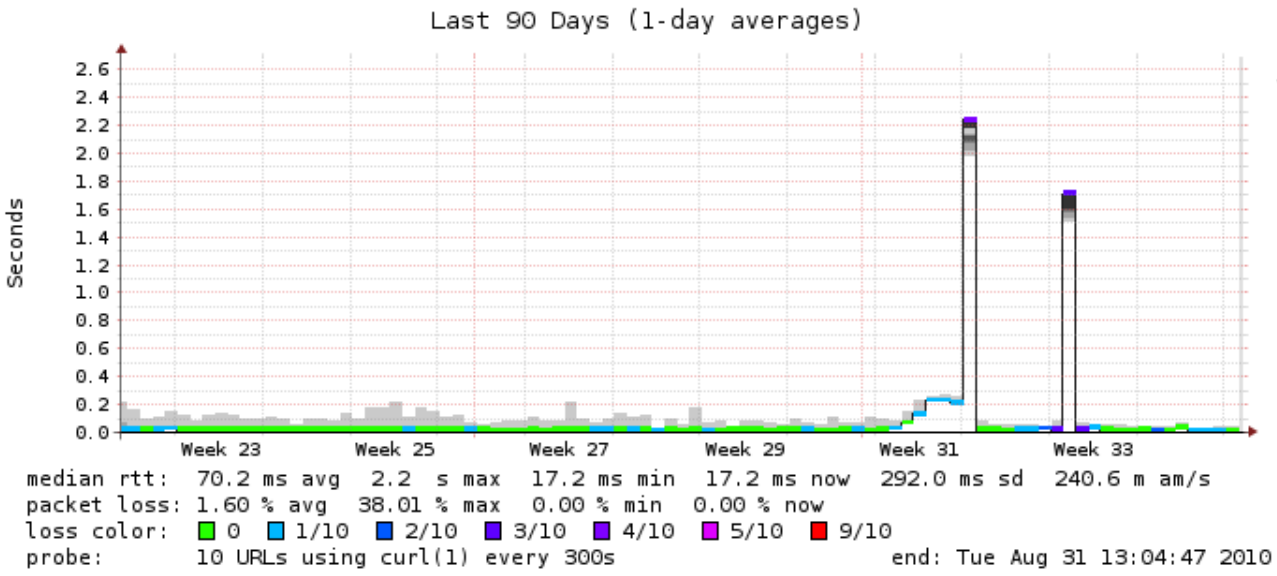


Figure 2 SmokePing Monitoring Results for TasksList target for 90 days

In weeks 31 and 33 there were some performance issues with this target, which we have tracked down to a new feature, which was introduced at that time. Consequently, this new feature has been optimized to perform better and response time has returned to normal for this target.

We also have a baseline monitoring, which measures network latency and packet loss and allows us to quantify better the web service performance metrics, in an independent way from network performance. This baseline monitoring is illustrated in Figure 3.

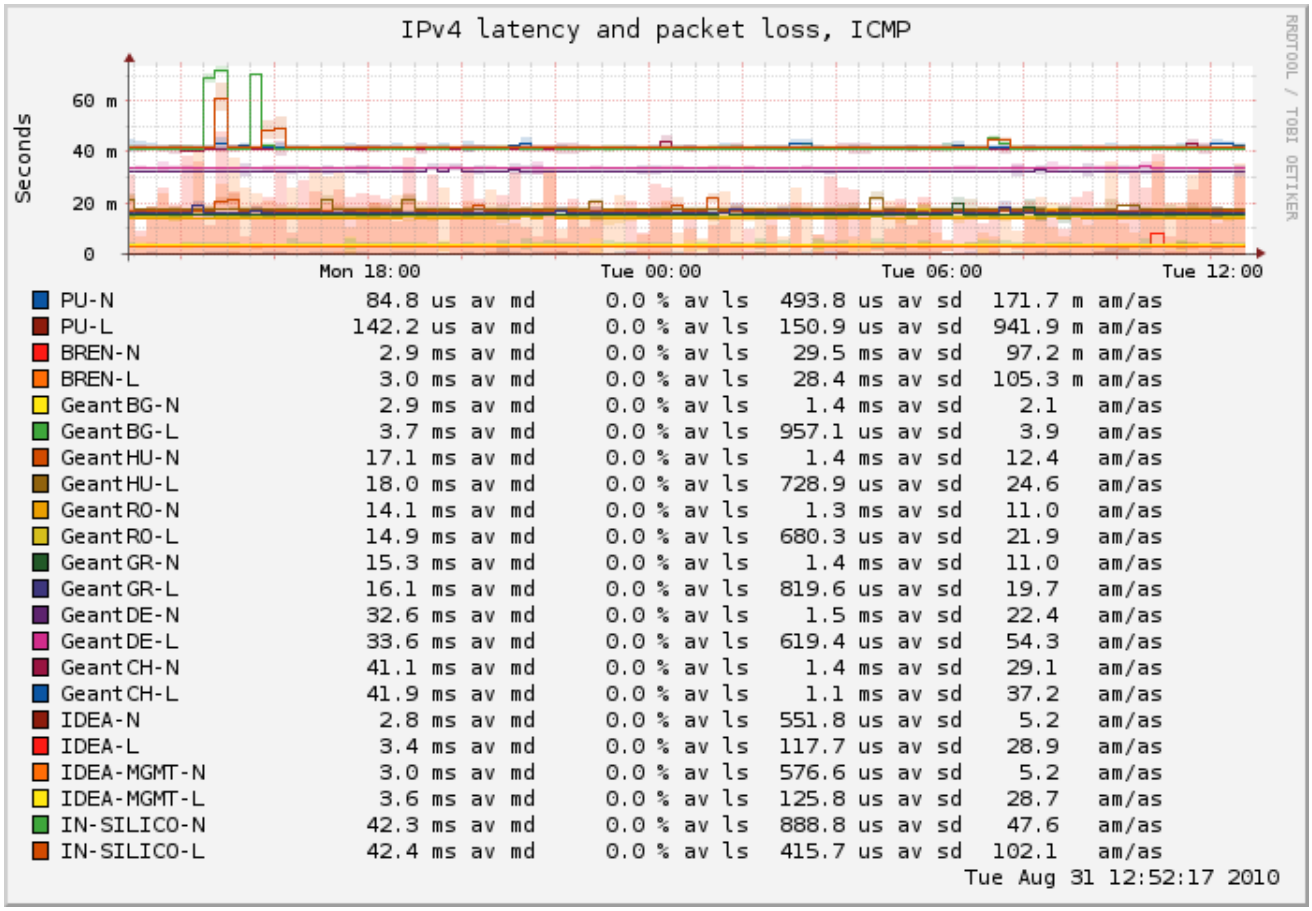


Figure 3 SmokePing Baseline Monitoring Results for OpenTox services

In particular, it is important to note the low network latency and the absence of packet loss, which highlights the fact that any latency/loss issues observed during that time are inherent to the monitored web services.

We intend to include more OpenTox web services in this automated and continuous availability and performance monitoring, in order to help developers and/or maintainers to find possible low performance culprits and to optimize their services. Another direction for further work is to find ways to better quantify the OpenTox web services regarding the remaining software quality metrics, which are not subject to the SmokePing monitoring (reliability, usability, maintainability, portability and interoperability). The best approach would be to seek an independent review for these metrics.

3. Improved API and Interface Designs

OpenTox is currently interacting with "external developers" i.e., developers involved in the development of other software applications or databases. For example, we are currently collaborating on the integration of the Leadscope database containing FDA *in vivo* data and the Titanium software for mining of human adverse events, so as to create new integrated OpenTox resources, accompanied by authorization and authentication. Another activity that has commenced in 2010 is work by external developers to get other major platforms (e.g., Bioclipse, CDK) interoperating with OpenTox for predictive toxicology purposes. Feedback from these interactions in addition to beta testing results provided much useful insight into the OpenTox API, including demonstration of its successes and feasibility, and areas for extensions and improvements. As a result an updated OpenTox API 1.2 was developed throughout 2010, and which will be frozen towards the end of 2010.

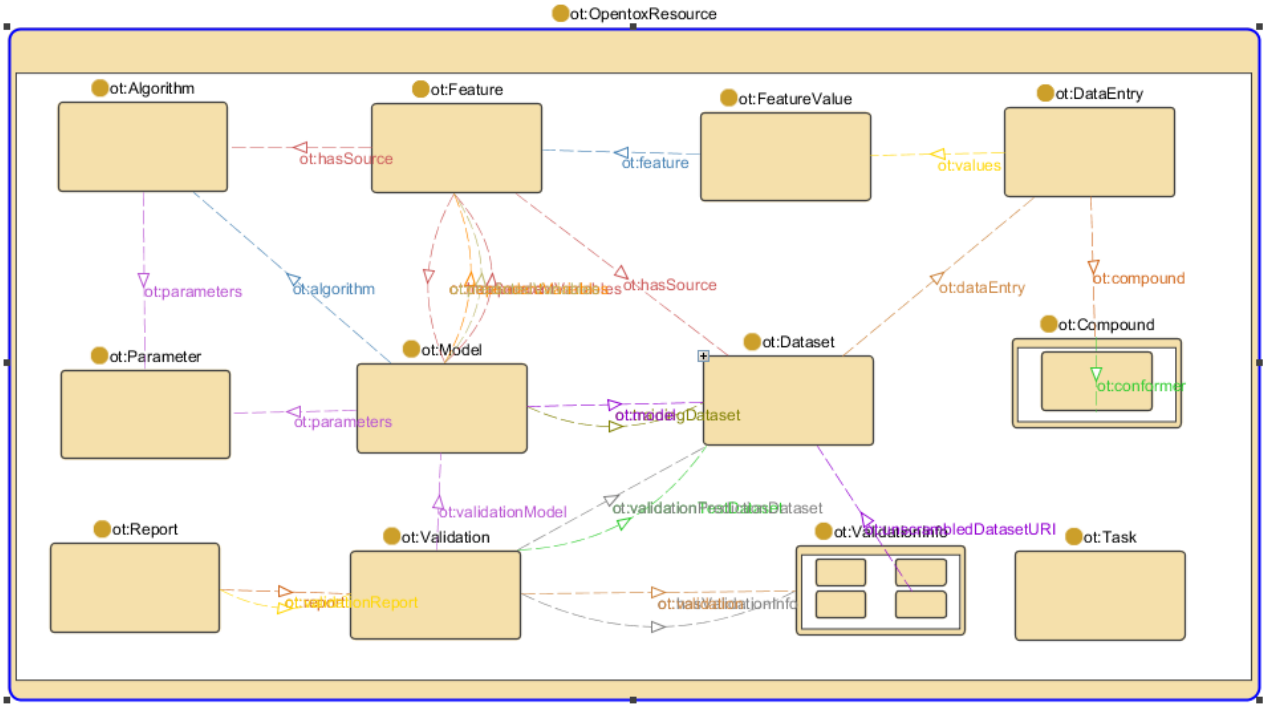
Introduction to OpenTox API 1.2

This section provides an overview to the proposed OpenTox API version 1.2. OpenTox components are currently web services with a REST (en.wikipedia.org/wiki/Representational_State_Transfer) interface. OpenTox interfaces are described at opentox.org/dev/apis/ and comprise:

- Compound
- Feature
- Dataset
- Algorithm
- Model
- Validation
- Task
- Ontology
- Authentication and Authorisation

Overview

The OpenTox ontology (opentox.org/api/1.1/opentox.owl) models the OpenTox web services as objects in RDF/XML. Relationships between OpenTox resources are modelled in the OpenTox ontology. The image is generated by the Protégé ontology development and editing software.


ot:OpentoxResource
ot:Algorithm
ot:Feature
ot:FeatureValue
ot:DataEntry
ot:hasSource
ot:feature
ot:values
ot:parameters
ot:algorithm
ot:hasSource
ot:dataEntry
ot:compound
ot:Compound
ot:conformer
ot:Parameter
ot:Model
ot:Dataset
ot:parameters
ot:predictedDataset
ot:validationModel
ot:Report
ot:Validation
ot:Task
ot:reportValidation

Common specifications for all OpenTox APIs

Parameters

Parameters are posted with a "Content-Type:application/x-www-form-urlencoded" HTTP header. Parameter names are typed in **bold** letters in the API definitions. Square brackets (e.g. **compound_uris[]**) indicate that a list of arguments is expected. We do not list all default arguments here. For a complete specification, see the online API documentation (opentox.org/dev/apis/api-1.2/AA).

For curl POST requests the -d/--data option should be used to ensure the content type. See the curl website for more information on the -d parameter.

Example:

curl -X GET http://{server}/dataset?compound_uris[]={compound_uri1}&compound_uris[]={compound_uri2})

curl -X POST -d 'dataset_uri=http://{server}/dataset/5' http://{server}/algorithm/xxx

Request and submit formats

The default OpenTox format is RDF/XML (with exception of the compound API), but service developers may support additional formats. You can request them, by specifying the MIME type in the "Accept" and "Content-Type" HTTP headers. We do not list all possible MIME types here. For a complete specification, see the online API documentation (opentox.org/dev/apis/api-1.2/AA)

If the service cannot serve the requested format, the default format (usually RDF/XML) will be returned.

Examples:

Request a compound in SDF format:

curl -X GET -H "Accept:chemical/x-mdl-sdfile" http://{server}/compound/{id}

Submit a compound in InChI format:

curl -X POST -H "Content-Type:chemical/x-inchi" --data-binary "InChI=1S/C5H10/c1-2-4-5-3-1/h1-5H2" http://{server}/compound

Create a new dataset:

curl -X POST -H "Content-Type:application/rdf+xml" --data-binary@my_data_file.rdf http://{server}/dataset

File uploads

Files can be uploaded by specifying "multipart/form-data" in the Content-Type header.

HTTP status codes

The following table gives some HTTP status codes that are used within OpenTox:

Interpretation	Nr	Name
Success	200	OK
Processing (for tasks)	202	Accepted
Unauthorised Access	401	Unauthorised
Resource not found	404	Not Found
Incorrect content	400	Bad request
Internal Server Error	500	Internal Server Error
Service not available	503	Service unavailable

REST API

Compound

Provides different representations for chemical compounds with a unique and defined chemical structure.

Description	Method	URI	Parameters	Result	Status codes
Search for compounds	GET	/compound	**search** **sameas** (=URI_FROM_ONTOLOGY) **tokenid**	List of compounds matching the query	200,404,503
Get the representation of a compound	GET	/compound/{id}	**feature_uris[]** (opt.) **tokenid**	Compound representation in one of the supported MIME formats, if **feature_uris[]** provided includes features and values	200,404,503
Create a new compound	POST	/compound	Compound representation in a supported MIME format **tokenid**	URIs for new compounds	200,400,503
Update a compound (opt.)	PUT	/compound/{id}	Compound representation in a supported MIME format **tokenid**	-	200,400,404,503
Delete a compound (opt.)	DELETE	/compound/{id}	-	-	200,400,404,503

Features per Compound

Description	Method	URI	Parameter	Result	Status codes
Get available feature URIs for a compound	GET	/compound/{cid}/feature	**feature_uris[]** **tokenid**	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** **value** **tokenid**	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** **value** **tokenid**		200,400,404,503
Delete specified features from the compound	DELETE	/compound/{cid}/feature	**feature_uris[]** (opt.) **tokenid**		200,400,404,503

Conformers

Optional support for multiple (e.g. 3D) structures per chemical compound (single structure by default)

Description	Method	URI	Parameters	Result	Status codes
Get available structures of a compound	GET	/compound/{id1}/conformer/	**tokenid**	List of structure URIs	200,404,503
Create a new structure	POST	/compound/{id1}/conformer	**tokenid**	New URI /compound/{id1}/conformer/{id2}	200,400,503
Remove all structures	DELETE	/compound/{id1}/conformer/	-	-	200,400,404,503
Get the representation of a structure	GET	/compound/{id1}/conformer/{id2}	**feature_uris[]** **tokenid**	Representation in a supported MIME format , with feature values , if **feature_uris**[] provided	200,404,503
Update the representation of a structure	PUT	/compound/{id1}/conformer/{id2}	**tokenid**	URI /compound/{id1}/conformer/{id2}	200,400,404,503
Remove a structure	DELETE	/compound/{id1}/conformer/{id2}	**tokenid**	-	200,400,404,503

Features per Conformer

Description	Method	URI	Parameter	Result	Status codes
Get available feature URIs for a compound	GET	/compound/{cid}/conformer/{cid}/feature	**feature_uris[]** **tokenid**	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}/conformer/{cid}/feature	**feature_uri** **value** **tokenid**	URI of the compound with the new feature, e.g. /compound/{id}/conformer/{cid}?feature_uris[]=the-new-feature	200,400,503
Update a new feature value	PUT	/compound/{cid}/conformer/{cid}/feature	**feature_uri** **value** **tokenid**		200,400,404,503
Delete specified features from the compound	DELETE	/compound/{cid}/conformer/{cid}/feature	**feature_uris[]** (opt.) **tokenid**		200,400,404,503

Feature

A Feature is an object, representing any kind of property, assigned to a Compound. The feature types are determined via their links to ontologies (Feature ontologies, Descriptor ontologies, Endpoints ontologies).

Description	Method	URI	Parameters	Result	Status codes
get description of a specific feature definition	GET	/feature/{id}	**tokenid**=token	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 **tokenid**=token	URI of the new feature definition	200,400,404, 503
update feature	PUT	/feature/{id}	**Content-type**="any-of-RDF-types", content=RDF-representation **tokenid**=token	-	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 **tokenid**	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

Dataset

Provides access to chemical compounds and their features (e.g. structural, physical-chemical, biological, toxicological properties).

Description	Method	URI	Parameters	Result	Status codes
Get a list of available datasets	GET	/dataset	query parameters (opt.) **tokenid**	List of URIs or RDF for the metadata only	200,404,503
Get a dataset	GET	/dataset/{id}	**tokenid**	Representation of the dataset in a supported MIME type	200,404,503
Query a dataset	GET	/dataset/{id}	**compound_uris[]** and/or **feature_uris[]** **tokenid**	Representation of the query result in a supported MIME type	200,404,503
Get metadata for a dataset	GET	/dataset/{id}/metadata	**tokenid**	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	**tokenid**	List of compound URIs	200,404,503
Get a list of all features in a dataset	GET	/dataset/{id}/features	**tokenid**	RDF or List of feature URIs (pointing to feature definitions/ontologies)	200,404,503
Create a new dataset	POST	/dataset	Dataset representation in a supported MIME type via **Content-type** **tokenid**	New URI /dataset/{id} or redirect to task URI (for large uploads)	200,202,400,503
Update a dataset	PUT	/dataset/{id}	Dataset representation in a supported MIME type via **Content-type** **tokenid**	Dataset URI or task URI	200,202,400,404,503
Remove a dataset	DELETE	/dataset/{id}	**tokenid**	-	200,404,503
Remove a part of the dataset	DELETE	/dataset/{id}	**compound_uris[]** and/or **feature_uris[]** **tokenid**	-	200,404,503

Algorithm

Provides access to OpenTox algorithms.

Description	Method	URI	Parameters	Result	Status codes
Get URIs of all available algorithms	GET	/algorithm	**sameas**=URI-OF-owl:sameAs-ENTRY (opt.) **tokenid**	List of all algorithm URIs	200,404,503
Get the ontology representation of an algorithm	GET	/algorithm/{id}	-	Algorithm representation in one of the supported MIME types	200,404,503
Apply the algorithm	POST	/algorithm/{id}	**dataset_uri** **prediction_feature** **dataset_service** **tokenid** parameters (opt.)	*model URI* *dataset URI* *featureURI* Redirect to task URI for time consuming computations	

Model

Provides different representations for QSAR/toxicology models.

Description	Method	URI	Parameters	Result	Status codes
Get a list of all available models	GET	/model	**query**=uri (opt.) **tokenid**	List of model URIs	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_uri** **dataset_service_uri** **tokenid**	URI of result dataset or task URI if	200,202,400,404,500,503
Apply a model to predict a compound	POST	/model/{id}	**compound_uri** **tokenid**	Prediction in a supported MIME type or task URI	200,202,400,404,500,503

Validation

A validation corresponds to the validation of a model on a test dataset. The results are stored in another dataset. Parameters with default values are optional.

Description	Method	URI	Parameters	Result	Status codes
Get all validations	GET	/	**tokenid**	List of validation URIs	200,404
Retrieves a validation representation	GET	/{id}	**tokenid**	Validation representation in one of the supported MIME types	200,404
Validates a model on a test dataset	POST	/	**model_uri** **test_dataset_uri** **test_target_dataset_uri** **tokenid**	Validation URI or Task URI	200,400,404,500
Builds a model on a training dataset and validates it on a test dataset	POST	/	**algorithm_uri** **prediction_feature** **algorithm_params** **training_dataset_uri** **test_dataset_uri** **test_target_dataset_uri** **y_scramble** **y_scramble_seed** **tokenid**	Validation URI or Task URI	200,400,404,500
Splits a dataset into training and test dataset according to a certain ratio, and performs a validation	POST	/training_test_split	**algorithm_uri** **prediction_feature** **algorithm_params** **dataset_uri** **split_ratio** **random_seed** **y_scramble** **y_scramble_seed** **tokenid**	Validation URI or Task URI	200,400,404,500
Deletes a validation.	DELETE	/{id}	**tokenid**	-	200,404

Task

Asynchronous jobs are handled via an intermediate Task resource. A resource, submitting an asynchronous job should return the URI of the task.

Description	Method	URI	Parameters	Result	Status codes
Get a list of all available tasks	GET	/task	**query**=task_status **tokenid**	List of URIs text/uri-list RDF representation application/rdf+xml	200,503,401
Get the representation of a running task	GET	/task/{id}	**tokenid**	Task representation in application/rdf+xml	200,201,202,404,503,401
Delete (cancel) a task	DELETE	/task/{id}	**tokenid**		200, 404, 503,401

Ontology

Provides storage and search functionality for objects, defined in OpenTox services and relevant ontologies.

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 the query results	200, 404, 500
Predefined query to retrieve all endpoints	GET	/ontology/endpoints		RDF representation of all endpoints	200, 404, 500
Predefined query to retrieve all algorithms	GET	/ontology/algorithms		RDF representation of all algorithms	200, 404, 500
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 the RDF representation and adds it to the RDF storage, thus making it available for the	200, 404, 500, 502

				subsequent queries. Any non-empty subset of parameters is valid (i.e. only query, only model_uri, query and algorithm_uri, etc.)	

Authentication and Authorisation

Granting access to protected or confidential resources for authorised users is handled by Authentication and Authorization (A&A).

Description	Method	URI	Parameters	Result	Status codes
Authentication	POST	/auth /authenticate	**username** **password** **uri**	token (Valid)	200, 401
Token validation	POST	/auth /isTokenValid	**tokenid**	Boolean	200
Logout	POST	/auth/logout	**tokenid**	void	200
Authorization	POST	/auth/authorize	**uri** **action** **tokenid**	Boolean (Grant/Deny) Boolean (Deny)	200, 401

Policies

Description	Method	URI	Parameters	Result	Status codes
Create a policy	POST	/pol	Policy representation in application specific XML format. **tokenid**=token		200, 400, 500
List policies	GET	/pol	**tokenid**=token	List of policy names	200, 500
List policy id	GET	/pol	**tokenid**=token **id**=id	XML representation of policy id	200, 401, 500
GET owner of URI uri	GET	/pol	**tokenid**=token **uri**=uri	Owner of policy id	200, 401, 500
Delete policy id	DELETE	/pol	**tokenid**=token **id**=id		200, 400, 401, 500

4. Conclusions

The initial OpenTox application prototype implementations ToxCreate and ToxPredict were evaluated through beta testing. Additionally, input from external developers, was obtained with regards to their use of the OpenTox API 1.1. An approach to automated continuous monitoring was implemented and deployed, helping to ensure the development of high performance, scalability and efficiency in OpenTox web services. Use cases guiding development continued to be developed and reviewed in parallel. The activities in combination were fed into discussions and the development of an improved API 1.2, including the new authorisation and authentication service for the integration of confidential data.

Appendix A ToxCreate Beta Testing Evaluation Template

In this section we give instructions for beta testers for ToxCreate (www.toxcreate.org) as well as a sample filled-in beta test form.

General Instructions

Please complete the ToxCreate Beta Test Tasks described below. To run the ToxCreate software you would need a web browser (a recent version of Firefox or Internet Explorer) and a network connection to Internet. Please answer the questions on the attached form, either by hard copy, or by editing an electronic copy of this document. Please return your feedback to Vedrin Jeliazkov vedrin.jeliazkov@gmail.com. With your permission, we may contact you occasionally during the course of the beta testing to solicit interim feedback. You might also want to register at the OpenTox site[1] and provide further feedback through the test case development issue tracker[2].

The ToxCreate software implements a prototype use case of the OpenTox framework, which enables end users to build new models for a selected endpoint and training set. The main steps of the workflow are listed below:

1. Input endpoint and training set;
2. Train and validate a model;
3. Use the model for predictions.

Beta Testing Objectives

The main objectives of this beta testing exercise are:

- To evaluate ToxCreate's technical capabilities and scientific value;
- To evaluate ToxCreate's ease of use and interactivity;
- To evaluate the end user documentation;
- To identify errors/bugs;
- To compile and prioritise a wish list of missing features, to be implemented in subsequent versions of the OpenTox framework.

Beta Testing Tasks

1. Complete Error! Reference source not found. (provide your name and contact details, web browser type/version and time period when the testing has been performed).
2. Open the following URL in your web browser http://toxcreate.org/
3. Proceed with functional evaluation of ToxCreate by following as many variants of the provided workflow as possible. These activities aim to evaluate the software's basic ability to generate the expected results, in the way you need them. Report your findings in Error! Reference source not found..
4. Complete Error! Reference source not found.. This section asks you to rate various aspects of the software using a 5-point scale.
5. List any bugs or problems in Error! Reference source not found. as you proceed.
6. Please answer any other relevant questions listed in Error! Reference source not found..

---

[1] www.opentox.org/join_form

[2] www.opentox.org/dev/testing/testcasedevelopment/testcasedevelopmentissuetracker

Part-A: Identification

Your Name	David Gallagher
Your Organisation	DG
Your Phone number	503 830 2772
Your E-mail address	gallagher.da@gmail.com
Used web browser (type/version)	Chrome 5.0.375.127
Time period when the testing has been performed	23rd August 2010

Part-B: Functional Evaluation

Test Case ID	Function	Tested? (yes/no)	Comments, Ideas and Issues
ToxCreate-01	Input a dataset	Yes	Failed to load a dataset, initially. 1. I clicked "Choose File" 2. I navigated to the test file "EPAFHM.csv" provided, then clicked "Open" 3. The name of the file appeared to the right of (Choose File" box) 4. I then clicked "Create model" 5. Error message returned "Please enter an endpoint name" 6. Started again, this time added "endpoint-name", then file loaded successfully. Suggestion 1: It may be more user-friendly if the file can be loaded without adding any name first, then the user is presented with a list of possible end-point names from those found in the file, or can create a new one.

Test Case ID	Function	Tested? (yes/no)	Comments, Ideas and Issues
ToxCreate-02	Train a model	Yes	This step (model training) started immediately the file was loaded and there seems to be no way to stop or abort the calculation. Also, there is no indication of how long the calculation will take. Suggestion 2: Before the calculation starts, let the user view the fields in the file and select an endpoint or create a new one. Then allow them to "Start" or "cancel". Also, while the calculation is running, provide a "Cancel" button. Suggestion 3: The user has no indication of how long the calculation will take (seconds or months?), so provide a progress bar with some status information.
ToxCreate-03	Inspect the model	Yes/No	Calculation did not end during the testing period, so I "inspected" a previously run model. Suggestion 4: The report is very cryptic and would benefit from added explanation including the significance of the various fields. Suggestion 5: It seems confusing to see many other models already computed, as I expected to see only mine. I suggest that other models are hidden by default or called up separately.

Test Case ID	Function	Tested? (yes/no)	Comments, Ideas and Issues
ToxCreate-04	Make predictions	yes	The "Predict" function has been implemented in a very easy to use way with the minimum of steps. Nice! A few minor issues were found: I drew in a structure (TNT) then clicked the Predict button (but forgot to select an endpoint). The structure disappeared and the error message appeared "Please enter a compound identifier and select an endpoint from the list". Suggestion 6: Correct the error message i.e. "Please select an endpoint from the list" and do not delete the drawn structure. Also, all endpoints should be selected by default. I am not sure which model was used for the prediction? Was it the model I had created, or another model? How can I tell? Suggestion 7: Can the details of the model be clearly identified (i.e. who created it and when with validation information)?

Part-C: Overall Comments and Usability Evaluation

Usability Question	Rating Scale 1 – Strongly Disagree 2 – Somewhat Disagree 3 – Neither Agree, Nor Disagree (No Opinion) 4 – Somewhat Agree 5 – Strongly Agree	Specific Comments on Rating
Overall		
This software is useful to me now, or it will be in the near future	N/A	
System output and visualization are useful and meet my needs	3	Needs status information such as a Progress bar during calculations
Software has the capabilities I need (note any exceptions)	3	Limited QSAR functions currently. Need "Cancel" function to stop calculations.
General impression is good (why?)	4	Clear and easy to read style
Software was easy to apply to my specific situation	N/A	
Data entry effort is manageable	3	Limited options for data entry. SD file format is also important.

Usability Question	Rating Scale 1 – Strongly Disagree 2 – Somewhat Disagree 3 – Neither Agree, Nor Disagree (No Opinion) 4 – Somewhat Agree 5 – Strongly Agree	Specific Comments on Rating
Technical Content		
Appropriate technical and scientific basis is used		unknown
Uses proper terminology	3	Report is too cryptic or technical to be understood by anyone except an expert
Performs calculations correctly		Not tested
Toolbars, menus, commands and options are appropriate	4	Need “Cancel” function to stop calculations.
Labels and terms are accurate and easy to understand (if not, what would you prefer?)	4	
Data formats are useful (if not, what would you prefer?)	3	SD is an important data format (isn’t this an a simple file translation issue, such as with Babel?)
I entered my own data and received the expected results	Not tested	
Boundary values (largest and smallest chemical samples) were handled correctly	Not tested	
Software Operation		
Trouble-free operation	4	Slow to respond, for example when clicking “Inspect”
Easy to navigate within the software	4	
Consistent and logical flow in using the software	4	
Easy to find what you are looking for	4	Did not know how to find and use the model I had created for prediction.
Software works as expected (uses standard user interface features)	4	Yes, in most cases (see some exceptions above)
Software works well within its family of software applications (if known)		
Files import and export to other needed applications	?	Can my model be imported into ToxPredict for my use?
Prints properly to a printer	4	Some minor formatting problems when printing
Documentation		
Clearly describes software purpose	Not tested	
Organization is clear and logical	Not tested	

Usability Question	Rating Scale 1 – Strongly Disagree 2 – Somewhat Disagree 3 – Neither Agree, Nor Disagree (No Opinion) 4 – Somewhat Agree 5 – Strongly Agree	Specific Comments on Rating
Examples show how to use the main features (please list any features needing more explanation or examples)	Not tested	
Tables, graphs & figures provide sufficient guidance through major software options	Not tested	
Do error messages clearly direct the user to a solution?	Not tested	
On-line help: was it easy to find what you wanted?	Not tested	
Included necessary technical support information	Not tested	
Appearance		
Colours, symbols, and graphics are legible and pleasing	5	Very nice clean and uncluttered layout
Looks professional	4	1. Positioning of some items and fonts could be improved. For example, headers are usually left justified because this is most intuitive for people who read from left to right. However, the most important header "ToxCreate" is right-justified, which demotes its significance and relevance. This means the eyes are first drawn to the "Create | Inspect | Predict" tabs which may be confusing. 2. Smaller fonts for the descriptive text would make it compete less with the tab names. 3. The inclusion of the "Tag line" is very helpful to provide a quick overview of the purpose of ToxCreate. The impact could be further improved by including a key benefit such as, reducing animal testing: e.g. "creates models to predict toxicity, reducing animal testing"
Correct spelling & grammar	5	
Application windows have consistent look and feel	4	

Part-D: Specific Bugs and Problems Noted

Please use the issue tracker at http://github.com/helma/opentox-toxmodel/issues to

1. report bugs (this should include a description of the problem and a list of steps to reproduce the problem). If you want to provide datasets, screenshots, etc you can send them directly to helma@in-silico.ch
2. suggest new features;
3. make general comments;
4. vote to prioritize bugfixes and feature additions;

If this is too much hassle, you can also use the form below.

Test Case ID (e.g. ToxCreate-01, ToxCreate-02, ..., ToxCreate-xy)	Nature of Problem	Full List of Steps to Reproduce the Problem
ToxCreate-dg01	When inputting dataset, it would be more user-friendly if the file can be loaded without adding any end-point name first. Then the user is presented with a list of possible end-point names from those found in the file, or can create a new one.	See Part B above
ToxCreate-dg02	Cancel option needed: Before a calculation starts, let the user view the fields in the file and select an endpoint or create a new one. Then allow them to "Start" or "cancel". Also, while the calculation is running, provide a "Cancel" button.	See Part B above
ToxCreate-dg03	Progress bar needed: The user has no indication of how long the calculation will take (seconds or months?), so provide a progress bar with some status information.	See Part B above
ToxCreate-dg04	The report is rather cryptic and would benefit from added explanation including the significance of the various fields. (Presumably, this will be taken care of when the reporting services are ready)	See Part B above
ToxCreate-dg05	"Inspect" page: too much data? It seems confusing to see many other models already computed, as I expected to see only mine. I suggest that other models are hidden by default or called up separately.	See Part B above
ToxCreate-dg06	Predict drawing tool: I drew in a structure (TNT) then clicked the Predict button (but forgot to select an endpoint). The structure disappeared and the error message appeared "Please enter a compound	See Part B above

	identifier and select an endpoint from the list". Suggest correcting the error message i.e. "Please select an endpoint from the list" and do not delete the drawn structure. Also, all endpoints should be selected by default.	
ToxCreate-dg07	Which model is being used for the "Prediction", and where is "my model": I am not sure which model was used for the prediction? Was it the model I had created, or another model? How can I tell? Can the details of the model(s) be clearly identified (i.e. who created it and when with validation information)?	See Part B above

Part-E: Other Generic Topics

Please comment on the following (if relevant):

- scientific value of algorithms included

  not evaluated

- speed of user interface interactivity and of calculations

  seems slow, up to 10 seconds to respond to a click (Inspect), a progress bar with status info would be helpful to indicate that the system is still working and not crashed

- order of screens and steps, and number of steps to complete an action

  seems good, though not fully tested yet

- compatibility of the software with existing workflows

  not fully consistent with ToxPredict

- organization of menu items

  Good

- quality of written explanations

  good, but more detailed information would be helpful

- terms or abbreviations used

  in the term "LAZAR" helpful or necessary on the front page, assuming ToxCreate will include more than just LAZAR?

- annoying or frustrating experiences

  slow response and no feedback when clicking on tabs

Appendix B ToxPredict Beta Testing Evaluation Template

In this section we give instructions for beta testers for ToxPredict (www.toxpredict.org) as well as a sample filled-in beta test form.

General Instructions

Please complete the ToxPredict Beta Test Tasks described below. To run the ToxPredict software you would need a web browser (a recent version of Firefox or Internet Explorer) and a network connection to Internet. Please answer the questions on the attached form, either by hard copy, or by editing an electronic copy of this document. Please return your feedback to Vedrin Jeliazkov (vedrin.jeliazkov -(at)- gmail.com). With your permission, we may contact you occasionally during the course of the beta testing to solicit interim feedback. You might also want to register at the OpenTox site[3] and provide further feedback through the test case development issue tracker[4].

The ToxPredict software implements a prototype use case of the OpenTox framework, which enables end users to run existing endpoint-specific models on a given compound (or dataset) and get model predictions. The main steps of the workflow are as follows:

1. Select input compound (enter chemical name, registry identifier (e.g. CAS, EINECS), SMILES, InChI, arbitrary keyword, SMARTS or draw molecule in molecular editor);
2. Select specific endpoint (e.g. Human Health Effects / Carcinogenicity);
3. Select one or more models, available for this particular endpoint (e.g. ToxTree: Benigni/Bossa rules for carcinogenicity and mutagenicity);
4. Apply selected model(s);
5. View and/or retrieve the resulting report, available in various formats, e.g. HTML, SDF, CML, SMI, PDF, XLS, ARFF or RDF.

Beta Testing Objectives

The main objectives of this beta testing exercise are:

- To evaluate ToxPredict' technical capabilities and scientific value;
- To evaluate ToxPredict' ease of use and interactivity;
- To evaluate the end user documentation;
- To identify errors/bugs;
- To compile and prioritise a wish list of missing features, to be implemented in subsequent versions of the OpenTox framework.

Beta Testing Tasks

7. Complete Error! Reference source not found. (provide your name and contact details, web browser type/version and time period when the testing has been performed).
8. Open the following URL in your web browser http://toxpredict.org
9. Proceed with functional evaluation of ToxPredict by following as many variants of the provided workflow as possible. These activities aim to evaluate the software's basic ability to generate the expected results, in the way you need them. Report your findings in Error! Reference source not found..

[3] www.opentox.org/join_form

[4] www.opentox.org/dev/testing/testcasedevelopment/testcasedevelopmentissuetracker

10. Complete Error! Reference source not found.. This section asks you to rate various aspects of the software using a 5-point scale.
11. List any bugs or problems in Error! Reference source not found. as you proceed.
12. Please answer any other relevant questions listed in Error! Reference source not found..

Known ToxPredict Problems

1. Bugs/usability problems:
    a. The overall GUI design is subject to improvement.
2. Missing features:
    a. The integrated online help is under development;
    b. Support for batch processing of datasets is under development;
    c. Support for file upload is under development;
    d. Support for molecular structure drawing is under development;
    e. Support for SMARTS searching is under development;
    f. Model integration is under development (only ToxTree, pKa and selected TUM (TU München) models are fully supported at the time of this writing);
    g. Models are available only for a subset of endpoints.

Part-A: Identification

Your Name	Roman Affentranger
Your Organisation	Douglas Connect
Your Phone number	
Your E-mail address	roman@douglasconnect.com
Used web browser (type/version)	Firefox v3.6.2
Time period when the testing has been performed	March 31, 2010

Part-B: Functional Evaluation

Test Case ID	Function	Tested? (yes/no)	Comments, Ideas and Issues
ToxPredict -01	Input chemical structure as SMILES	Yes	Is the interpretation of the input SMILES string case sensitive? CCCCOC and ccccoc both are interpreted as 1-methoxybutane, even though according to http://inchi.info/converter_en.html the latter should be (1E)-1-methoxybuta-1,3-diene (the search criterion was set to "equal"). Not sure if ccccoc even is a valid SMILES at all, though. The issue seems to be worth checking, however. C1CCCCC1 and c1ccccc1 are, however, interpreted correctly as cyclohexane and benzene, respectively. It would help to make sense of these results if the SMILES string of the found hit would be displayed.
ToxPredict -02	Input chemical structure as MOL	Yes	Loaded a mol file for SMILES C4C2OCOC2CCC4CN([H])C(=O)CN(C1=O)NCN3C1CC(N3)C5CCCCC5, worked fine. However, when browsing for files, the filter is set to "all files". It would be nice if it would filter for the allowed file types. Curiously, while uploading the said molecule as MOL worked fine, entering its SMILES wouldn't yield a hit. When entering a SMILES string, and selecting "Equal" for the search criterion, the software shouldn't be going to look up the entry in a database, but instead just draw the molecule.

Test Case ID	Function	Tested? (yes/no)	Comments, Ideas and Issues
ToxPredict -03	Input chemical structure as SDF	No	
ToxPredict -04	Input chemical structure as InChI	Yes	Generated InChi strings for CCCCOC on http://inchi.info/converter_en.html (resulting in InChI=1S/C5H12O/c1-3-4-5-6-2/h3-5H2,1-2H3). While a SMILES search for CCCCOC yielded methoxybutane, entering the corresponding InChi string yielded no hit.
ToxPredict -05	Input chemical structure as compound name	Yes	Entering 1-methoxybutane gave no hits. Entering e.g. "hexane" and setting the search criterion to "equal" results in many hits merely containing the string "hexane", but not in hexane itself. Only searching for "Hexane" (capital "H") yields hexane. Not even "hexane" with search criterion "Sounds like" finds hexane.
ToxPredict -06	Input chemical structure as CAS number	No	
ToxPredict -07	Input chemical structure as EINECS number	No	
ToxPredict -08	Input chemical structure as SMARTS	No	
ToxPredict -09	Input chemical structure as arbitrary string	No	
ToxPredict -10	Input chemical structure through the integrated molecular structure editor	Yes	Similar to entering a SMILES and searching for "Equals", it doesn't make sense that no hits are found when drawing a structure and selecting "Search for Structure", while at the same time the drawn molecule is accepted when uploading it as a MOL file.
ToxPredict -11	Verify selected structure(s)	Yes	There's not much to verify here, as selection of individual hits among the presented ones isn't possible yet. If one needs to go back to step 1 for some reason, all previous entry is lost (e.g. the drawing, search string, etc.), and one has to start again from scratch. It would help a lot if that information would still be there, to allow the user to modify the previous search string, drawing, etc. When presenting the hits for validation, it would be helpful to present/summarize the search information at the top of the display, e.g., present the search string, and say what it has been interpreted as.

Test Case ID	Function	Tested? (yes/no)	Comments, Ideas and Issues
ToxPredict -12	Select a relevant model from a list of available models	Yes	I found it difficult to make sense of the presented models. There should be information for each of them. Also, clicking on the entries under "algorithm" only yields an xml without associated style. Clicking on the "YES" under "Descriptors" yields a download. Of what? The file downloaded for the model "Predictive QSAR Model generated by the algorithm mlr" (named "independent") just contains 10 lines similar to http://apps.ideaconsult.net:8180/ambit2/feature/21926. There are two models that look exactly the same (4th and 5th models, called "Predictive QSAR Model generated by the algorithm mlr"). They are probably different, though (see next point). The download of the training datasets is a bit weird. When trying to download the SDF file for the 4th model ("Predictive QSAR Model generated by the algorithm mlr") one ends up on the website http://apps.ideaconsult.net:8180/ambit2/dataset/20?max=20?media=chemical%2Fx-mdl-sdfile, while for the 5th model (which is indistinguishable from the 4th one on the model selection page, one downloads a file called "269" (without a file extension). For the sixth model ("OpenTox model created with TUM's kNNregression model learning web service"), the SDF file is called "23". Some more sensible filenames would be great, and also extensions would help a lot… Turns out that for the other download formats, the behaviour is similar. For the 4th model, one always ends up on the above-mentioned website. For the 5th and 6th models, one is prompted a download of files 269 and 23, respectively. The behaviour is somewhat different when clicking on the pdf format: instead of prompting to save the file, it is displayed in the browser, but not in a new window…

Test Case ID	Function	Tested? (yes/no)	Comments, Ideas and Issues
ToxPredict -13	Apply model(s) (make a prediction(s))	Yes	Test performed for the loaded mol file SMILES: C4C2OCOC2CCC4CN([H])C(=O)CN(C1=O)NCN3C1CC(N3)C5CCCCC5 - Molecular weight: ok - CPSA: all values are NaN - Zagreb Index: ok - Predictive QSAR Model generated by the algorithm mlr (4th model): It would process for about 30 seconds, then the browser prompts: "To display this page, Firefox must send information that will repeat any action (such as a search or order confirmation) that was performed earlier.", offering the buttons "Resend" and "Cancel". After having resent a couple of times without finishing, I suspect that the computations restart every time… - Next "Predictive QSAR Model generated by the algorithm mlr (5th model): after processing for a few seconds, I got "Bad Request (400) - Bad Request" next to the green check mark that usually indicates that the computation is finished. - OpenTox model created with TUM's kNNregression model learning web service: similar to above, but the error message here was "Internal Server Error (500) - Internal Server Error✔" - ToxTree: Benigni/Bossa rules for carcinogenicity and mutagenicity: works fine, but the results are a bit weird: structural alerts = NO, Potential mutagen/carcinogen = NO, but unlikely to be mutagen/carcinogen is also NO. -pKA, all ToxTree, Lipinski, XlogP: ok
ToxPredict -14	Follow the progress of a prediction task	Yes	ok
ToxPredict -15	View predictions and experimental data (HTML format)	Yes	There might be a bit more explanations on what the results mean… Either I didn't come across any, or it isn't clear which are experimental data, and which are predictions.

Test Case ID	Function	Tested? (yes/no)	Comments, Ideas and Issues
ToxPredict -16	Retrieve resulting report in SDF format	Yes	Works, however, the file that is downloaded is lacking the .sdf extension. Also, it does not contain any results, but only the 2D-molecules
ToxPredict -17	Retrieve resulting report in CML format	Yes	Not offered. I could only download in SDF format.
ToxPredict -18	Retrieve resulting report in SMI format	Yes	Not offered. I could only download in SDF format.
ToxPredict -19	Retrieve resulting report in PDF format	Yes	Not offered. I could only download in SDF format.
ToxPredict -20	Retrieve resulting report in CSV format	Yes	Not offered. I could only download in SDF format.
ToxPredict -21	Retrieve resulting report in ARFF format	Yes	Not offered. I could only download in SDF format.
ToxPredict -22	Retrieve resulting report in RDF format	Yes	Not offered. I could only download in SDF format.

Part-C: Overall Comments and Usability Evaluation

Usability Question	Rating Scale 1 - Strongly Disagree 2 - Somewhat Disagree 3 - Neither Agree, Nor Disagree (No Opinion) 4 - Somewhat Agree 5 - Strongly Agree	Specific Comments on Rating
Overall		
This software is useful to me now, or it will be in the near future	4	
System output and visualization are useful and meet my needs	3	Explanations of the results are missing
Software has the capabilities I need (note any exceptions)	4	
General impression is good (why?)	4	
Software was easy to apply to my specific situation	4	
Data entry effort is manageable	3	Whenever moving back to step 1, all previous information is lost (e.g., one has to re-draw the molecule, or re-type the SMILES, etc.).
Technical Content		
Appropriate technical and scientific basis is used	4	

Usability Question	Rating Scale 1 – Strongly Disagree 2 – Somewhat Disagree 3 – Neither Agree, Nor Disagree (No Opinion) 4 – Somewhat Agree 5 – Strongly Agree	Specific Comments on Rating
Uses proper terminology	4	
Performs calculations correctly	3	Not sure about CPSA and Benigni/Bossa
Toolbars, menus, commands and options are appropriate	4	"Next" (or even the whole flowchart) could be replicated at the bottom of the screen.
Labels and terms are accurate and easy to understand (if not, what would you prefer?)	4	
Data formats are useful (if not, what would you prefer?)	4	Mol2 and pdb format for structure upload would be nice
I entered my own data and received the expected results	-	-
Boundary values (largest and smallest chemical samples) were handled correctly	-	-
Software Operation		
Trouble-free operation	2	Frequently received error codes 400 and 500, however in a non-reproducible way. That is, navigating again to toxpredict.net and repeating the steps sometimes passed the point where I received these errors before.
Easy to navigate within the software	4	
Consistent and logical flow in using the software	4	
Easy to find what you are looking for	3	There should be more background information on the models.
Software works as expected (uses standard user interface features)	4	
Software works well within its family of software applications (if known)	-	-
Files import and export to other needed applications	-	-
Prints properly to a printer	-	-
Documentation		
Clearly describes software purpose	-	Not available
Organization is clear and logical	4	4
Examples show how to use the main features (please list any features needing more explanation or examples)	-	Not available

Usability Question	Rating Scale 1 – Strongly Disagree 2 – Somewhat Disagree 3 – Neither Agree, Nor Disagree (No Opinion) 4 – Somewhat Agree 5 – Strongly Agree	Specific Comments on Rating
Tables, graphs & figures provide sufficient guidance through major software options	-	Not available
Do error messages clearly direct the user to a solution?	2	
On-line help: was it easy to find what you wanted?	-	Not available
Included necessary technical support information	-	Not available
Appearance		
Colours, symbols, and graphics are legible and pleasing	3	Individual compounds should be better separated in the results table.
Looks professional	4	
Correct spelling & grammar	3	Didn't check for spelling and grammar
Application windows have consistent look and feel	4	

Part-D: Specific Bugs and Problems Noted

Test Case ID (e.g. ToxPredict-01, ToxPredict-02, ..., ToxPredict-xy)	Nature of Problem	Full List of Steps to Reproduce the Problem
TP-RA-01	SMILES sometimes not case sensitive???	ccccoc with "Equal" search -> presents CCCCOC as hit.
TP-RA-02	CPSA yiels all NaN	Enter CCO, select as only model "CPSA descriptor"
TP-RA-03	Benigni/Bossa results are confusing	Enter CCO, select as only model the Benigni/Bossa model: Structural Alert for genotoxic carcinogenicity NO Structural Alert for nongenotoxic carcinogenicity NO Potential S. typhimurium TA100 mutagen based on QSAR NO Unlikely to be a S. typhimurium TA100 mutagen based on QSAR NO Potential carcinogen based on QSAR NO Unlikely to be a carcinogen based on QSAR NO

		For a better assessment a QSAR calculation could be applied. NO Negative for genotoxic carcinogenicity YES Negative for nongenotoxic carcinogenicity YES

Part-E: Other Generic Topics

Please comment on the following (if relevant):

- scientific value of algorithms included
- speed of user interface interactivity and of calculations
- order of screens and steps, and number of steps to complete an action
- compatibility of the software with existing workflows
- organization of menu items
- quality of written explanations
- terms or abbreviations used
- annoying or frustrating experiences

  The SMILES of the found hits are not displayed in step 2. Neither is the search string and the search options. To verify the results, it would be nice to have a summary of the input at the top: one could then see immediately that one selected “Equal” instead of “Sounds like”, etc.

When drawing a molecule and not finding any hits, the drawing is lost when going back to step 1. Maybe one would just like to remove a group from the drawn molecule, and not have to re-draw the whole thing…

Uploading files is a little unintuitive…