OpenTox

Deliverable D1.4

Evaluation of the Final Implementation

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

SEVENTH FRAMEWORK PROGRAMME

Contract No.	Health-F5-2008-200787
Document Type:	Deliverable Report
WP/Task:	WP1 / D1.4
Name	Evaluation of the Final Implementation
Document ID:	OpenTox Deliverable Report WP1-D1.4
Date:	October 30, 2011
Status:	Final Version
Organisation:	DG, DC

Contributors		
	David Gallagher	DG
	Barry Hardy	DC
	Roman Affentranger	DC

Distribution:	Public

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

Document History:	1 – First Draft Aug 30, 2011
	2 – Second Draft Oct 26, 2011
	3 – Third Draft Oct 31, 2011

Table of Contents

Table of Figures

Summary

This report summarises the testing and evaluation carried out on OpenTox in the final year of the project. The evaluation included assessment of OpenTox from the following points of view a) framework, b) standards, c) ontology, d) data resources, e) services, and f) applications.

OpenTox reached a major milestone in Spring 2010 with the development of its first two prototype applications ToxPedict and ToxCreate, which were successfully demonstrated live at a workshop ahead of the first AXLR8 meeting in Potsdam. These applications have continued to be improved and were successfully used in the final hands-on OpenTox workshop in Munich in August 2011 where they received overall very positive evaluations by a variety of users. Additional application developments in the third year of the project have included the development of the Bioclipse-OpenTox integration, the QPRF reporting editor Q-Edit, the CheS-Mapper visualisation tool, and the Taverna workflow integration whose evaluation is reported in this report. These applications have demonstrated that a variety of Graphical User Interfaces (GUIs) may build upon the OpenTox foundation of resources to provide a diverse set of experiences to users through user-friendly interfaces and environments, significantly increasing the user community for OpenTox. The Bioclipse and Taverna integrations are particularly significant as they demonstrate the successful integration of OpenTox with two other major platforms, thus accelerating the capabilities and value creation of OpenTox. These developments all build and take advantage of the significant investment in the design of the framework and its open standards carried out during the research project.

In 2011 an OpenTox Virtual Machine was developed. This allows the user to download and install a virtual machine on their computer where OpenTox applications can be run in a local environment. Hence although OpenTox supports a distributed set of resources, it may also be collapsed to run as a local stand-alone application. We brought the virtual machine software to SETAC Africa and successfully installed it on ca. 40 laptops of workshop participants.

Online tutorials have also been created supporting usage of the applications on an ongoing basis. International use in 2011 has been broadened and demonstrated by the coordinator who for example has worked with risk assessors at Health Canada and African scientists at SETAC Africa in the use of the applications in workshops, in addition to responding to a number of presentation invitations from industry and international conference organisers.

These developments were disseminated and evaluated through a number of means, ranging from internal testing and evaluation by non-developers within the OpenTox project, to provision of online tutorials and exercises, to a hands-on workshop involving 9 exercises carried out in the final month of the project, which was subsequently made available online for further evaluation, and to dissemination at a variety of international conferences. This report summarises the results, feedback and evaluation of these developments of the OpenTox framework and its services and applications. We also provide recommendations for next steps and future directions.

Internal testing and evaluation supported the development of a number of OpenTox resources, services and applications so that they reached the stage of prototypes that could be provided to users to carry out exercises in a workshop setting or on their own. Underlying these is the OpenTox Application Programming Interface (API), which is an important output of the FP7 research project. It provides a specification against which development may continue by both OpenTox partners and the broader development community. In addition to further development of existing resources and applications, new ones may be created compliant with the specifications, thus supporting a growing set of interoperable linked resources for the field of toxicology. As the specification is an open standard it may also continue to develop and improve. As the latest version of the API (1.2) supports Authorisation and Authentication, such infrastructure can support the integrated use of both open and commercial resources. Policies for controlled access to resources can also be implemented which supports user access to resources based on licences, data sharing agreements and legal contracts. The use of OpenTox API-compliant REST services to communicate instructions between linked resources with URI addresses supports the use of a wide variety of commands to carry out operations such as data integration,

algorithm use, model building and validation. OpenTox thus provides a "lower level programming language" to the more technically-oriented user for carrying our research, executing tasks or application development.

The OpenTox ontology development made good progress in the last year of the project and included development of an ontology to support endpoints and organ effects, and the capability to convert from the existing ToxML data standard into a semantic web format. From the toxicology ontology workshop held at the EBI resulted a perspectives review and a toxicology ontology roadmap to provide guidance to the field. As OpenTox creates a semantic web for toxicology, it should be an ideal framework for incorporating such future ontology developments, thus supporting both a mechanistic framework for toxicology in addition to best practices in statistical analysis, a combination that the field will surely profit from.

A number of (Q)SAR models were built with OpenTox for a number of REACH endpoints (carcinogenicity, mutagenicity, aquatic/fish toxicity, LogP), and a representative subset is presented in the report. Several of these models have been made available with validation reports through the ToxPredict application. We propose a "challenge" should be organised in which the community could develop different models using OpenTox services applied to different REACH endpoints.

The OpenTox 2011 conference in Munich, held at the end of the FP7 project in August 2011, attracted a diverse and knowledgeable audience including many principle investigators in the field, working in different areas. Hence, the group of ca. 80 scientists was strongly cross-disciplinary, cross-sector and international. With this backdrop we also ran a final workshop in which 9 different OpenTox exercises and applications were run by participants in groups in a hands-on format. It is difficult to communicate the intensity of the day as many participants engaged in a variety of activities all of which had been made possible by the project. This report summarises those experiences extended through additional experiences that went before and were continued afterwards through testing and evaluation and through the provisions of online tutorials.

The evaluation of OpenTox by these users provided a clear response that the project has produced very valuable and concrete results. Moreover, although still at the prototyping stage, many users were able to carry out activities with applications of significant scientific interest, including activities that had not been previously possible for them. The majority of evaluation answers to our questions produced answers in the excellent and very good categories, providing a sound endorsement by the user community that we have created something really valuable and worthwhile from the project. OpenTox should therefore be seen as a special success story the European Commission has supported through its research program.

At the end of every research project there are always issues that the work has surfaced which provide input into future work. Here we summarise some of the main issues which we currently have discovered or recognised:

1. OpenTox is a brand new framework, hence developers need to learn about the new approaches before being able to implement;

2. It is difficult to maintain the performance of a distributed set of resources and to ensure the ongoing evaluation and stability of core distributed services;

3. Certain users may prefer a local application, which is an area requiring future development by application developers;

4. Additional resources are needed to ensure the quality of a resource e.g., to curate a data source and to filter and label for quality;

5. Applications requiring integration of resources from multiple partners require a significant integrated testing methodology and effort;

6. A well-defined API is insufficient on its own to support developers in their development of resources to be accessed by other applications;

7. A significant effort is required to incorporate support for a broad range of datasets, protocols and models in the expanding field of predictive toxicology. Even more important is communication between different disciplines who do not understand each other very well;

8. The "21st century" mechanistic approach to predictive toxicology requires a framework and methods that place numerous demands beyond the building of a "simple" (Q)SAR for a single endpoint;

9. A substantial amount of work is needed to organise and harmonise the set of ontologies needed for predictive toxicology. This will take a significant coordinated effort and investment;

10. Further application development, communication skills and learning methodologies are needed to provide easy-to-use fit-for-purpose tools for end users not specialised in computer science techniques.

We provide here recommendations for the future development and sustainability of OpenTox and its benefit and impact:

1) OpenTox should be developed further to support creation of a global interoperable infrastructure for Predictive Toxicology to aid achievement of the goals of the 21st century vision. This could be supported through a dedicated EC infrastructure project supporting the goals of a number of programs including European activities under REACH, FP7 Health, FP7 Environment, EC JRC, and Innovative Medicines Initiative and international activities such as the US EPA ToxCast program and Tox21. The unique benefit of OpenTox lies in its core commitment to open standards which can support in a neutral way the goals of all stakeholders;

2) A program should be put in place to implement the Toxicology Ontology Roadmap. OpenTox can play a leadership role in collaboration with other stakeholders;

3) In consultation with industry, CEFIC and ECHA, an analysis should be carried out promptly to identify steps that could be taken to optimise OpenTox for supporting the activities of industry in meeting the 2013 REACH deadlines;

4) A business ecosystem approach providing value to industry and SMEs should be developed for OpenTox in support of its economic exploitation;

5) Consideration should be given to the formation of a legal structure for OpenTox e.g., as an international organisation or foundation;

6) The potential of OpenTox to be exploited in the development of safer medicines should be pursued;

7) OpenTox has evolved a community of interested parties beyond the initial FP7 partners. This should be developed further as an active and vibrant network;

8) OpenTox partners committed at their final project meeting to a combined interest and resolve to continue to develop OpenTox beyond the end of the project. Doing so will be rewarding!

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 (Hardy, Douglas et al. 2010).

We subsequently carried out an evaluation of the initial OpenTox APIs, the prototype applications and initial feedback from other development groups who had 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 updated API 1.2 (www.opentox.org/dev/apis/api-1.2) was extended to incorporate the lessons learned from these experiences. The API 1.2 was also updated to support services to secure confidential data within OpenTox. Specifically, Authentication and Authorization (A&A) interfaces were developed and implemented. We previously reported in 2010 on the evaluation of the framework and it's API in the OpenTox D1.3 report "Prototype evaluation, improved API, and interface designs". This report focuses on further developments and their evaluation that took place in the final year of the project.

In addition to the ToxPredict and ToxCreate prototype applications, a number of further application developments took place in the latter stages of the project:

a) The Bioclipse workbench was integrated with OpenTox to provide users the ability to combine local and distributed OpenTox models and to interactively explore changes in chemical features and their impact on toxicities;

b) Taverna workflows were implemented across OpenTox resources to support the execution of more complex tasks;

c) Support for the regulatory reporting of toxicity predictions was extended through the creation of a QPRF reporting application;

d) Metabolism prediction was incorporated based on SMARTCyp;

e) A Weight of Evidence (WoE) workflow was developed and applied to the holistic profiling and prioritisation of drug candidate libraries;

f) Three applications were developed supporting descriptor-based creation for model building: IBMCDesc, MaxTox, and ToxDesc;

g) Integration of commercial or non-open software and databases accompanied by security policies and authorisation and authentication was completed for 3 cases: Leadscope FDA toxicology databases; Pharmatrope adverse event modelling, and the Fraunhofer repeated dose toxicity database;

h) The CheS-Mapper application was developed supporting the interactive visualisation and exploration of datasets based on chemical features;

i) The OpenTox ontology was extended (as reported in the OpenTox D3.4 report on "Final database with additional content";

j) The validation and reporting service was developed for confidential datasets as supported by A&A services (as reported in the OpenTox D3.3 report "Tools for the integration of confidential data").

These developments were disseminated and evaluated through a number of means:

a) internal testing and evaluation by non-developers within the OpenTox project;

b) provision of online tutorials and exercises;

c) a hands-on workshop involving 9 exercises carried out in the final month of the project, which was subsequently made available online for further evaluation;

d) dissemination at a variety of international conferences.

This report summarises the results, feedback and evaluation of these developments of the OpenTox framework and its services and applications. Recommendations for next steps and future directions are provided.

2. Evaluation Methodology

2.1 Evaluation Criteria

The evaluation included an assessment of OpenTox from the following points of view a) framework, b) standards, c) ontology, d) data resources, e) services, and f) applications.

In addition to internal testing activities, many end users were involved in the evaluation and included a diversity of roles: by science, sector, activity, first-time users, and users from industry who advised us during the project.

The following main criteria for services and applications were adopted:
1. OpenTox API – interoperability, compliancy, benefit;
2. Use Case – completion of use case tasks, achievement of use case goals, variations;
3. Usability – design, user interface, user interaction, supporting information, help, tutorials;
4. Fitness for REACH purposes – ability to complete REACH-relevant tasks;
5. Technical Performance – reliability, robustness, scalability.

We also tried to take into account context differences e.g., an end user obtaining a prediction is quite different to an external application developer using or developing an OpenTox service.

2.2 Internal Testing and Evaluation

We report on a representative set of internal testing activities carried out on the following applications during the final stages of the project (depending on the results, many but not all of these testing activities were subsequently followed by external testing and evaluation):

a) **ToxPredict** (www.toxpredict.org)

ToxPredict provides a web-based application in which the user enters chemical structures and obtains information back from a registered and distributed set of OpenTox resources including the OpenTox database and registered models. The application informs the user what OpenTox currently knows about the toxicity-related properties of molecules.

b) **ToxCreate** (www.toxcreate.org)

ToxCreate takes a user-supplied dataset and builds a validated QSAR model, which may either be a classification or regression model. The current ToxCreate application is restricted to the use of the lazar algorithm only, although future extensions could include other algorithms.

c) **Taverna** workflows across OpenTox resources

We have used the Taverna workflow approach (www.taverna.org.uk) to build workflows across multiple OpenTox resources. Hence a user may in principle build any operation (e.g., data integration, model building, variation of parameters, profiling of a library, etc.) across multiple OpenTox resources for data, algorithms, model building, validation and reporting. The workflows may be saved for future use and adaptation. Hence the combination of Taverna and OpenTox potentially provides a powerful capability for scientific analysis.

d) **Bioclipse-OpenTox**

The successful integration by external developers Ola Spjuth and Egon Willighagen of the Bioclipse workbench with OpenTox was a major highlight of the third year of the project. It represented the first integration of another major platform with OpenTox, and provided a user-friendly interface to the remote running of OpenTox services.

e) **QPRF reporting**

Q-edit is a QPRF editor developed under OpenTox which generates a report on a (Q)SAR-based toxicity as recommended by the EC JRC for REACH reporting purposes. The Q-edit application supports the user in

documenting the results of the prediction based on OpenTox resources and additional expert information entered by the user.

f) SMARTCyp

SMARTCyp uses a combination of reference Quantum Chemical calculations and rules to predict in ranked order the primary metabolites of a parent compound. This capability has been integrated into OpenTox so that the user may examine such metabolic information or may potentially then build models based on metabolites rather than just the parent compound.

g) CheS-Mapper

CheS-Mapper is a visual application in which the user may interactively explore the features of an OpenTox dataset, and as a result gain insight into the contributions of the features to the toxicology properties of interest. The application may hence serve to provide insight, develop understanding and hypotheses, and aid in the interpretation of model results or suggest new experiments.

h) Validation

The OpenTox validation services may be used to apply a variety of different validation approaches to validate a predictive toxicology model created within the OpenTox framework. During Year 3 of the project this capability was extended to confidential datasets (for more details, see OpenTox D5.4 report on "Final Validation Routines with facilities to validate against confidential data")

2.3 User Testing and Evaluation

In the previous OpenTox D1.3 evaluation report, we reported user evaluation results based on beta testing templates which were developed for ToxPredict and ToxCreate. In the final testing and evaluation we created 9 exercises which are made available online at www.opentox.org/meet/opentox2011/exercises:

A. Use ToxPredict to obtain REACH-relevant information on compounds
David Gallagher, Barry Hardy (More Information ...)

In this exercise the user examines the toxicity information available on a number of compounds and then may try their own examples.

B. Use ToxCreate to build a (Q)SAR model
Christoph Helma, David Vorgrimmler (More Information ...)

In this exercise the user builds a (Q)SAR model and examines validation and confidence information on the model and its predictions

C. Validation and Reporting of Models
Andreas Maunz, Harry Sarimveis (More Information ...)

The user studies the use of OpenTox validation approaches and the validation information reported on a model.

D. Create and run Workflows across OpenTox services
Joerg Wicker, Fabian Buchwald (More Information ...)

In this exercise the user may try out Taverna-based workflows to integrate toxicology datasets and build a model.

E. Generate Predictions and their QPRF Reports
Pantelis Sopasakis, Sunil Chawla (More Information ...)

The Q-edit application is used to create a QPRF report based on OpenTox resources and a prediction.

F. Predict Metabolites of Compounds
Patrik Rydberg, Nina Jeliazkova (More Information ...)

The user can use the SMARTCyp approach to predict the metabolites of a compound.

G. Apply Bioclipse and OpenTox to modify Chemical Structure to remove Toxicity
    Ola Spjuth (More Information ...)

    Using a combination of Bioclipse and OpenTox the user may interactively study the toxicity properties of a molecule and how changes in chemical structure influence those properties.

H. Build a Weight of Evidence for Drug Candidate Molecules
    Roman Affentranger, Indira Ghosh (More Information ...)

    In this exercise we use OpenTox in a drug discovery context to prioritise anti-malarial drug candidates according to their predicted toxicities based on a Weight-of-Evidence approach.

I. Visualise and Explore Chemical Feature Space of a Dataset
    Martin Gütlein, Andreas Karwath (More Information ...)

    Users may explore the chemical features of datasets using an interactive visualization application.

These exercises were all subject to user evaluation as part of the final hands-on workshop in Munich (See OpenTox D6.5-8 Report on "Dissemination Summary Report for Year 3" for further details). These exercises were also subject to further evaluation through the online versions. Instructions for these exercises vary in terms of the amount of detail provided. We report here, as an example, the instructions for exercise A. on ToxPredict in Appendix D.

Evaluation Responses

The user evaluation reports for the above exercises is based on the survey provided at:

https://www.surveymonkey.com/s/652XY7C

3. Evaluation Results

3.1 Internal Testing and Evaluation

Appendix A reports on a representative set of internal testing activities carried out on the following applications during the final stages of the project. We summarise the results and actions here.

3.1.1 ToxPredict

Figure 1: ToxPredict – OpenTox Application for model prediction of toxicity endpoints for chemical structures involving the combination of distributed multiple predictive toxicology resources

As one of the first OpenTox prototypes, the ToxPredict application was subjected to a number of exposures and evaluations in 2010 and was used successfully in a workshop with risk assessors at Health Canada in early Spring 2011. The developers introduced a new interface in 2011 and added the capability of authorisation and authentication along with substantial extensions in functionality. However this new interface and the added functionality also introduced a significant change in the Graphical User Interface (GUI) and interaction with it. We hence carried out extensive internal testing of the new interface resulting in the reporting of numerous issues back to the developer. However our overall evaluation of the performance and directions of the application was positive. The resulting application was subsequently subject to numerous exposures to users in the Munich workshop and online.

3.1.2 ToxCreate

Figure 2: *ToxCreate – OpenTox Application for the creation and validation of a predictive toxicology model involving the combination of distributed multiple predictive toxicology resources*

The ToxCreate application did not change significantly in 2011; rather the focus was on addressing some of the reported issues with the initial prototype, and including authorization and authentication. Our internal testing found the application to be working well overall, although the following limitations were noted:

1) The application is currently only working with one algorithm, the lazar algorithm;

2) It is difficult to interpret the current approach to confidences in terms of their interpretation for Applicability Domain (AD);

3) Models created within ToxCreate when registered with the OpenTox ontology service and presented in ToxPredict often resulted in unsatisfactory performance e.g., in presentation to the user of unexplained results, validation and AD information.

3.1.3 Taverna

Results from the execution of the Taverna workflows were not found upon initial use to be so easy to understand and more extensive documentation was recommended. Also it was not so clear how parameters could be modified for modified workflow adaptation. Nevertheless the execution of tasks to accomplish an objective was impressively successful.

3.1.4 Bioclipse

The successful integration by external developers Ola Spjuth and Egon Willighagen of the Bioclipse workbench with OpenTox was a major highlight of the third year of the project. It represented the first integration of another major platform with OpenTox. Additionally, the initial testing of Bioclipse-OpenTox was very satisfactory, and showed a major advance in a user-friendly interaction with OpenTox resources. With our internal testing, new applications were run successfully within ten minutes of downloading and installation.
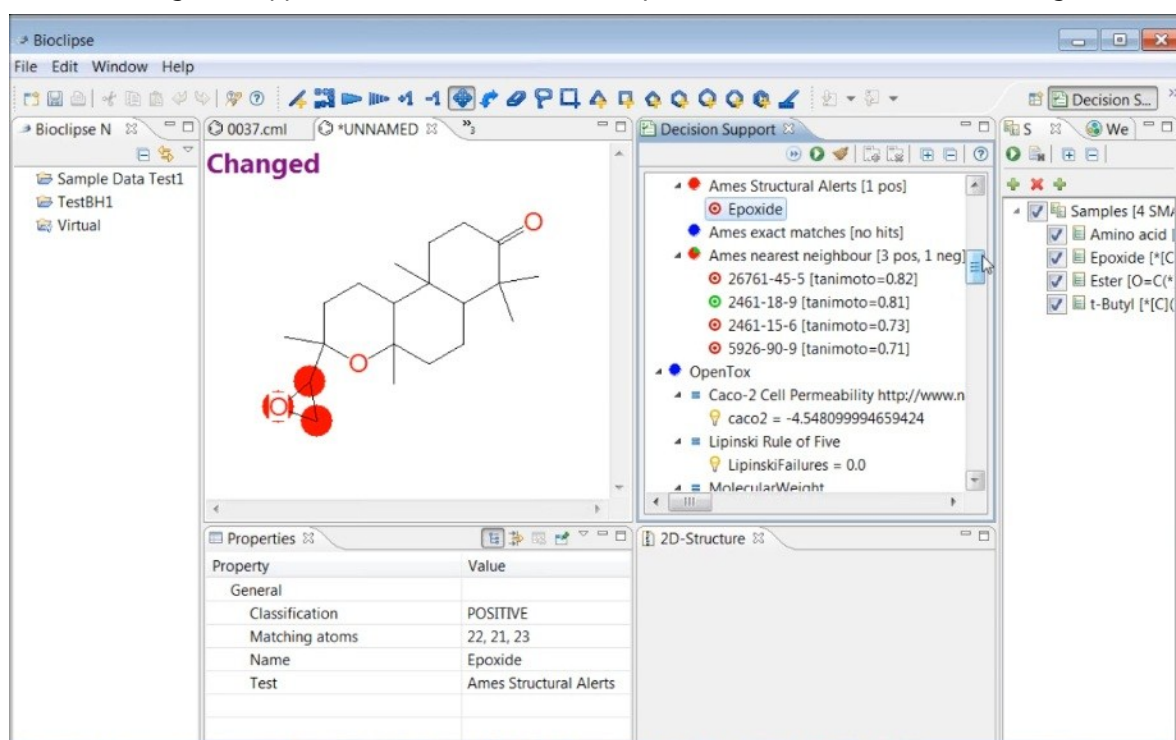
Figure 3: Bioclipse InterAction with OpenTox

3.1.5 QPRF reporting

Out initial testing of Q-edit showed it was difficult for the user to figure out what to do or how to populate the editor workflow with OpenTox information. This led to an improvement in documentation and guidance before exposing users to the new application for evaluation.
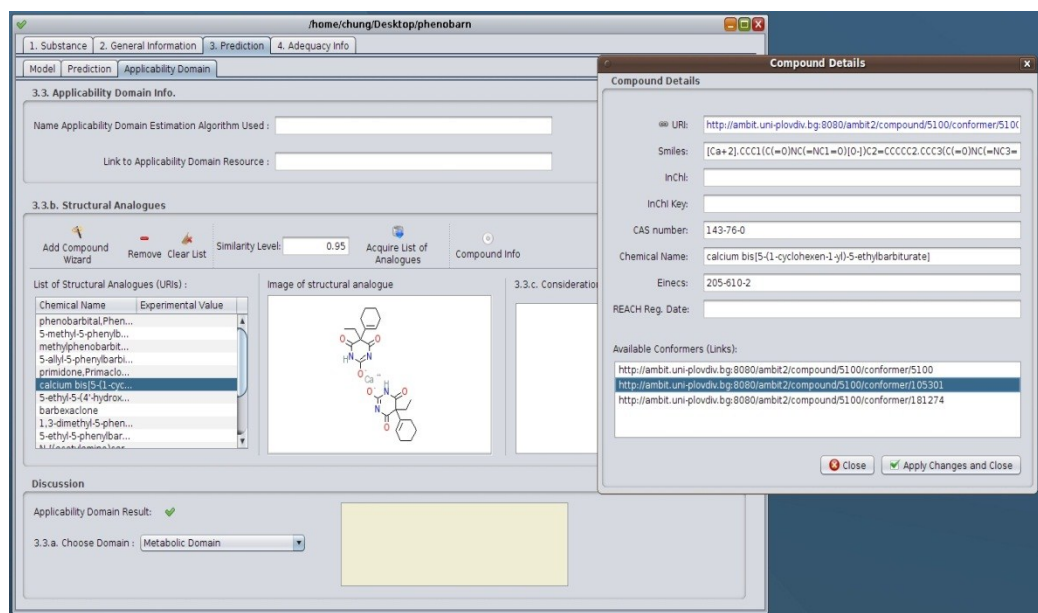
Figure 4: Creating a QPRF report with Q-edit

3.1.6 SMARTCyp

Initial testing of the SMARTCyp application revealed weaknesses in understanding of the information presented and the lack of explicit structures for primary predicted metabolites. These issues were subsequently addressed in preparation for presentation to users at the final OpenTox workshop in Munich.

3.1.7 CheS-Mapper

Initial testing showed that it was difficult to understand the options presented to the user for parameter selection. Also performance issues in dealing with compound sets of greater than a small number of molecules were encountered. These issues were addressed by the developers prior to the workshop in Munich, where the application was received positively and received a peer-reviewed best poster recognition.
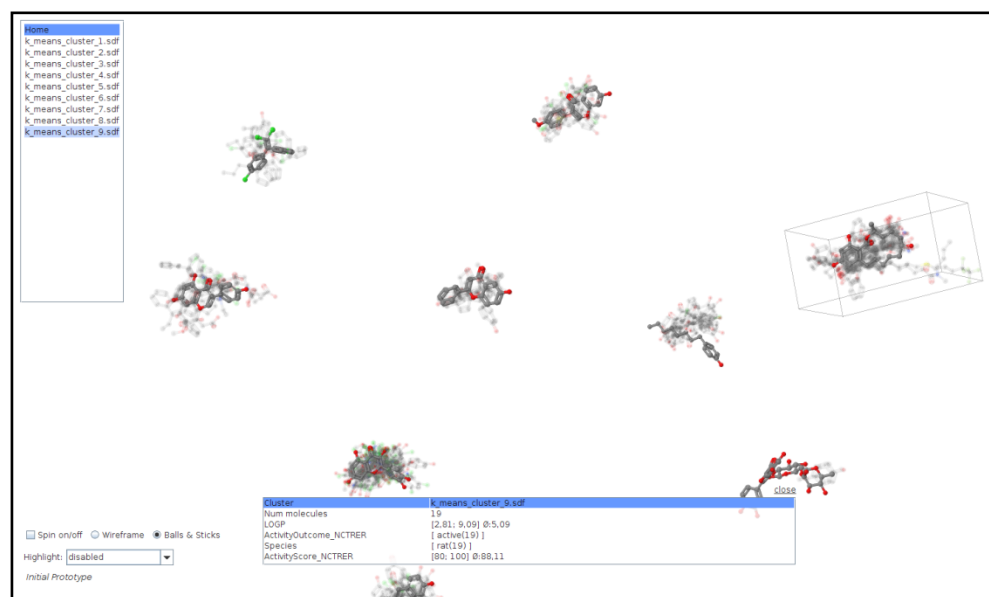
Figure 5: Visualisation of Clustered Compounds with CheS-Mapper

3.1.8 Validation

The validation services were found to work well if the user had the knowledge to use them. However the approach was found to be difficult for the non-insider to follow, and the developers responded with web interfaces to make the task easier to perform.
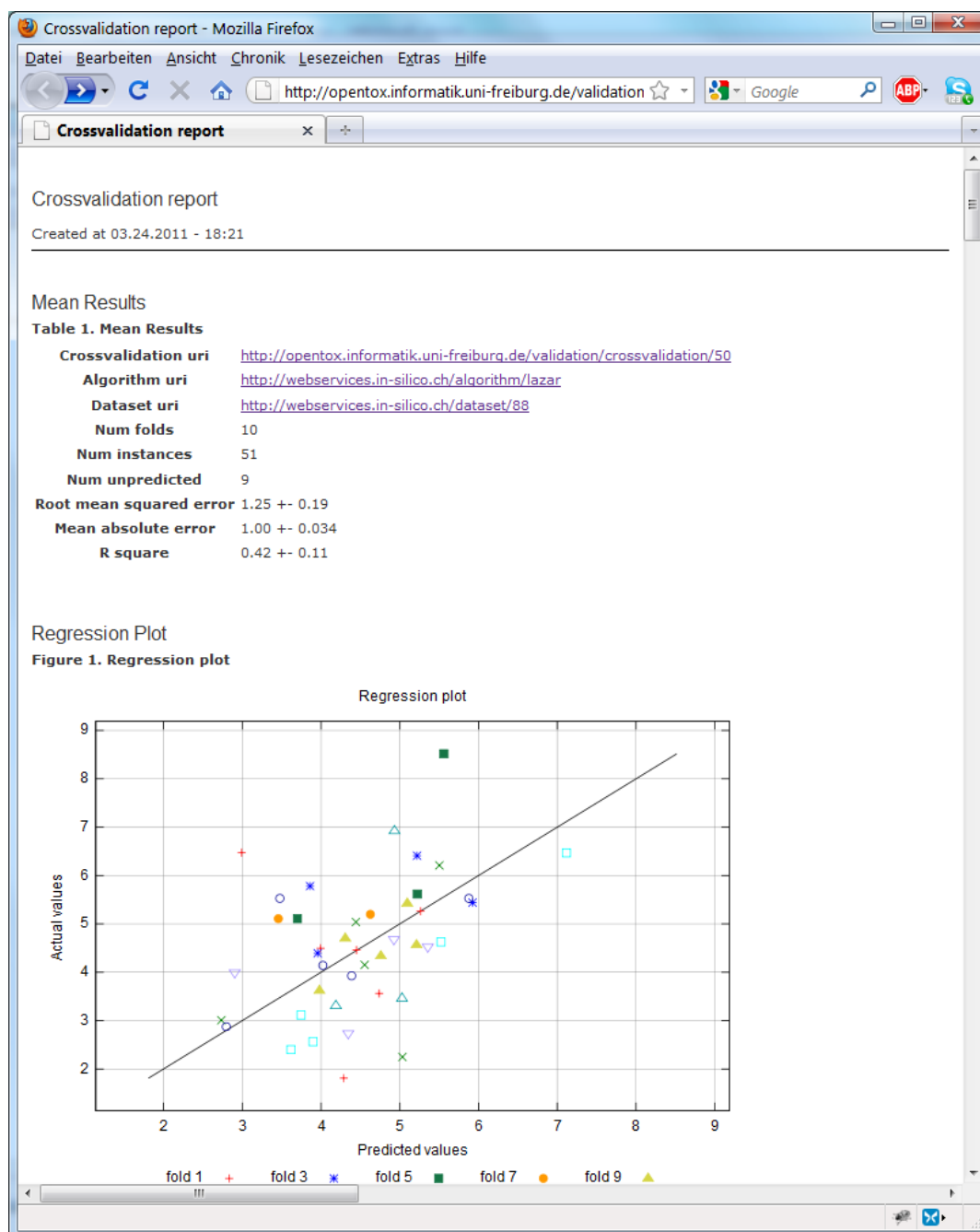
Figure 6: OpenTox Validation Report for a ToxCreate Model

3.2 OpenTox API

We provide the following narrative by an external developer who achieved the first major interoperation with OpenTox as a testimony:

"OpenTox has in my opinion taken many and large steps towards the important goal of an infrastructure for predictive toxicology. Taking its stance on open standards and open technologies, OpenTox has made it possible for us external parties to easily make use of the outcomes; both standards, technologies, data, and predictive models. It has been a pleasure interacting with the OpenTox community which has been very

responsive, inclusive, and productive. My hope is that this project will be developed further, and my opinion is that what has been developed so far provides an excellent starting point for this.

As the project leader of Bioclipse (www.bioclipse.net) we have been developing a project separate from OpenTox related to predictive toxicology, which provides execution and interpretation of toxicity models on the local computer. The addition of OpenTox services to this has meant that our project was capable of combining such local models with established models from the OpenTox community.

I will continue my research in the field of predictive toxicology and push for establishing a Swedish infrastructure for data and models related to this field. OpenTox will here have a major impact for my work, as the choice of standards and technologies will need to be interoperable with other European parties. I think that it will be very fruitful to continue the development of OpenTox into a sustainable and rich infrastructure for predictive toxicology, and I together with my Swedish partners would be very interested in being part of such continuation.

My group has used the OpenTox API for integrating access to services and data into Bioclipse. As the OpenTox API is built on standard technologies such as RDF and REST Services, this has been easy to integrate and provides extensive search, provisioning, and invocation functions. The importance of an ontology defining the concepts cannot be overstated. The A&A functionality is important as the first thing people remark on is that structures are sent over unprotected networks. We have not used the modeling, validation, or reporting tools. What could be improved is the results from predictive models to allow for interpreting results as e.g. chemical substructures that could be visualized. Also the framework is in some cases resulting HTML, which in my opinion should not be part of OpenTox but part of the presentation in any OpenTox client. Maybe some best practices when returning results from predictive models should be established."

3.3 OpenTox Services

Two of the exercises created for the OpenTox Munich workshop relied on direct use of OpenTox services. The summary of the evaluation results are provided in Appendix B. We review the experiences and feedback here. Answers to questions having a majority of excellent or very good are taken as particularly favourable, whereas a minority was taken as an indication of needed improvement.

3.3.1. Validation and Reporting of Models
Andreas Maunz, Harry Sarimveis

In this exercise users use a variety of web forms and cURL commands to run OpenTox validation and reporting web services and to examine their results. Models are built based on training set data created as an OpenTox data resource and subsequently validated using an external test set or cross-validation.

Users rated highly the scientific interest of the exercise, its technical merit, purpose and functionality, the understandability of the results and information generated, evaluation of the information provided, and information relevancy.

Users indicated average performance or a need for improvement in the ease of use, look and feel of the interface, system response to exploration, help and guidance.

Users found the flexibility of the approach, informative reports, use of resource addresses and transparency were attractive. Interest in applying the approach to human health risk assessment was expressed. Suggestions included adding confidence bounds and more results plots. Provision of further information for future use was requested.

3.3.2. Build a Weight of Evidence for Drug Candidate Molecules
Roman Affentranger, Indira Ghosh

In this exercise users apply a weight of evidence approach to examine the toxicity of anti-malarial drug candidate molecules. The user creates an OpenTox data resource for the molecules and then applies a range of

models to the dataset. The user may then examine the toxicity profiles of the molecules and prioritise candidates that have more attractive profiles.

Users rated highly the scientific interest of the software, technical merit, purpose and functionality, the understandability of the results and information generated, ease of understanding of what they were doing, system response to exploration, help and guidance, evaluation of the information provided, information reliability, information adequacy, and information completeness.

Users indicated average performance or a need for improvement in the ease of use, look and feel of the interface, and information relevancy.

Users liked the strength of the teaching and the instructions and guidance provided. More information on unfamiliar concepts such as features was requested.

3.4 OpenTox Applications

Seven of the exercises created for the Munich workshop relied on direct use of applications linked to OpenTox services. The summary of the evaluation results are provided in Appendix B. We review the experiences and feedback here. Answers to questions having a majority of excellent or very good are taken as particularly favourable, whereas a minority was taken as an indication of needed improvement.

3.4.1. Use ToxPredict to obtain REACH-relevant information on compounds
David Gallagher, Barry Hardy

ToxPredict provides a web-based application in which the user enters chemical structures and obtains information back from a registered and distributed set of OpenTox resources including the OpenTox database and registered models. The application informs the user what OpenTox currently knows about the toxicity-related properties of molecules. As of August 2011 ToxPredict included 24 validated models addressing 11 different endpoints.

Users rated highly the scientific interest of the software, ease of use, technical merit, look and feel of the interface, ease of understanding of what they were doing, system response to exploration, evaluation of the information provided, information reliability, and information adequacy.

Users indicated average performance or a need for improvement in purpose and functionality, the understandability of the results and information generated, help and guidance, evaluation of the information provided, information relevancy and information completeness.

Users liked the ability to easily obtain toxicity information on compounds, that all available information was available, and that the interface was simple and easy-to-use. Suggestions for improvement included requests for more explanations of the information provided to the user, clarification of the meaning of a Yes/No alert, explanation of confidence, provision of units, and more links to documentation and references. It was suggested that further work should continue to make the application more complete.

3.4.2 Use ToxCreate to build a (Q)SAR model
Christoph Helma, David Vorgrimmler

In this exercise users used ToxCreate to build (Q)SAR models, examined the results of models and their statistics, and then used the models to make predictions.

Users rated highly the scientific interest of the software, ease of use, technical merit, purpose and functionality, the understandability of the results and information generated, look and feel of the interface, ease of understanding of what they were doing, help and guidance, evaluation of the information provided, information reliability, information relevancy, information adequacy, and information completeness.

Users indicated average performance or a need for improvement in the system response to exploration.

Users liked the straightforward user interface, uniform design and the upload format flexibility.

3.4.3. Create and run Workflows across OpenTox services
Joerg Wicker, Fabian Buchwald

In this exercise the Taverna workflow approach (www.taverna.org.uk) is used to build workflows across multiple OpenTox resources. Examples for data integration and model building were worked on.

Users rated highly the scientific interest of the software, ease of use, look and feel of the interface, system response to exploration, information reliability, information relevancy, information adequacy, and information completeness.

Users indicated average performance or a need for improvement in the technical merit, purpose and functionality, the understandability of the results and information generated, ease of understanding of what they were doing, help and guidance, and evaluation of the information provided.

Users liked the ability to use the workflows to integrate different datasets.

3.4.4. Generate Predictions and their QPRF Reports
Pantelis Sopasakis, Sunil Chawla

In this exercise the Q-edit QPRF editor is used to generates a report on a (Q)SAR-based toxicity as recommended by the EC JRC for REACH reporting purposes.

Evaluation information from users is at this stage too preliminary to make conclusions.

3.4.5. Predict Metabolites of Compounds
Patrik Rydberg, Nina Jeliazkova

In this exercise the SMARTCyp method is used to predict the P450 sites of metabolism and to generate a list of ranked metabolites.

Users rated highly the scientific interest of the software and its ease of use. Users liked that the method worked on first principles and did not require experimental data for compounds. The method works well on several classes of compounds but there are compound classes where results need improvement.

3.4.6. Apply Bioclipse and OpenTox to modify Chemical Structure to remove Toxicity
Ola Spjuth

In this exercise the decision support functionality of Bioclipse is used to examine the toxicity properties of molecules including local and distributed OpenTox models. Users may interactively modify chemical structures and recalculate the toxicity predictions, thus supporting the interactive exploration of the effect of chemical modifications on toxicity properties.

Initial responses to the Bioclipse-OpenTox integration at a number of international workshops and conferences were very positive and included the OpenTox-EBI industry forum workshop on toxicology ontology, BIO-IT 2011, the SETAC Africa OpenTox workshop, and the final OpenTox Munich workshop.

Further evaluation information based on the online tutorial activity from users is at this stage too preliminary to make further conclusions.

3.4.7. Visualise and Explore Chemical Feature Space of a Dataset
Martin Gütlein, Andreas Karwath

In this exercise the user uses the CheS-Mapper visual application to interactively explore the features of an OpenTox dataset, and as a result gains insight into the contributions of the features to the toxicology properties of interest.

Users rated highly the scientific interest of the software, ease of use, technical merit, purpose and functionality, the understandability of the results and information generated, look and feel of the interface, ease of understanding of what they were doing, help and guidance, evaluation of the information provided, information reliability, information relevancy, and information completeness.

Users indicated average performance or a need for improvement in the system response to exploration and information adequacy.

Users appreciated the ability to visualize the chemical space of a set of molecules that is easily comprehensible and easy-to-use, and to allow identification of different features of subsets of molecules.

Users encountered errors in uploading datasets to the application indicating a need for further testing and bug fixing for this new application. Suggestions for further development included adding structural alerts and the ability to show overlap between different datasets.

4. Discussion

4.1 Testing and Evaluation Overview

Internal testing and evaluation supported the development of a number of OpenTox resources, services and applications so that they reached the stage of prototypes that could be provided to users to carry out exercises in a workshop setting and to provide applications and tutorials online for users to use on an ongoing basis. Common issues encountered in testing including service failure, unexpected system responses, lack of information, and GUI bugs. Application testing involving the integration of services from multiple partners often revealed further issues that resulted in unexpected or missing user experiences. For example one developer might create a validated model compliant with the OpenTox API, but registration of the service with the OpenTox ontology service did not ensure successful provision of the validation information to the user in the application. OpenTox as a research project involving the ongoing development of distributed resources by multiple partners provides a useful testbed for development. However the distributed and temporal nature of resources often resulted in lack of reliability in applications, e.g., OpenTox involves the ongoing creation of many resources with link addresses, but these links may be absent or failing upon future use. Hence, for future OpenTox-based infrastructure development and deployment, more controlled integrated testing and deployment processes will need to be invoked to guarantee quality of service levels and system reliability.

4.2 OpenTox API

The OpenTox API is an important output of the FP7 research project. It provides a specification against which development may continue by both OpenTox partners and the broader development community at large. In addition to further development of existing resources and applications, new ones may be created compliant with the specifications, thus supporting a growing set of interoperable linked resources for the field of toxicology. As the specification is an open standard it may also continue to develop and improve based on such future developments. For example, within the current ToxBank infrastructure project (www.toxbank.net) the OpenTox API is being extended to support inclusion of a variety of different protocols and datasets from the SEURAT-1 research program (www.seurat-1.eu) including biological protocols, *in vitro* assays, dose-response analysis, and integrated analysis across multiple omics datasets. The reuse of existing components in such infrastructure development accelerates the development of new services. As the latest version of the API (1.2) supports Authorisation and Authentication, such infrastructure can support the integrated use of both open and commercial resources. Policies for controlled access to resources can also be implemented which supports user access to resources based on licences, data sharing agreements and legal contracts.

4.3 OpenTox Services

The use of OpenTox API-compliant REST services to communicate instructions between linked resources with URI addresses supports the use of a wide variety of commands to carry out operations such as data integration, algorithm use, model building and validation. OpenTox thus provides a "lower level programming language" to the more technically-oriented user for carrying our research, executing tasks or application development. On the downside, commands and addresses can become quite verbose, so this "language" is not particularly elegant, concise or easy to understand. Hence only a limited number of core developers at this point are active in its use. Additionally, as OpenTox application development is a relatively new activity, more user friendly access to OpenTox capabilities for certain tasks, is currently not yet available to the broader community of users, who will not use such an approach. This situation should improve with further application development. Nevertheless, we have been able meanwhile to work through examples at OpenTox workshops in 2010 and 2011 with users to successfully carry out tasks such as the model validation exercise reported in this report. Documentation, guidance and tutorials and their ongoing development and maintenance, will therefore be a crucial component for OpenTox sustainability in the years to come.

4.4 OpenTox Databases

As of August 2011 the OpenTox database included quality-labelled data for 646,042 chemicals (including all chemicals currently registered under REACH). This valuable resource has been created by OpenTox partner Ideaconsult extending the Ambit database to be compliant with the OpenTox framework (see OpenTox D3.4 report "Final Database with additional content"; Jeliazkova, N and Jeliazkov, V. 2011). This is a critical resource for OpenTox as it provides data to developers for model building. It also provides useful toxicology information to users who for example need to gather all available information on a chemical as the first step in their REACH risk assessment preparation. Ambit itself is not a particularly user-friendly application for the general user; however this drawback has been addressed by the development of the ToxPredict application, where the general user may obtain information from the database through inputting a chemical name or drawing a structure (Hardy, Douglas et al 2010).

4.5 OpenTox Ontology

The OpenTox ontology development made good progress in the last year of the project and included development of ontology to support endpoints and organ effects. Additionally, the capability to convert from the existing ToxML data standard into a semantic web format was progressed. This work was presented at the Bio-Ontologies 2011 conference and was subsequently selected as an invited publication for peer review (Tcheremenskaia et al. 2011). In collaboration with the EBI Industry forum a toxicology ontology workshop was held at the EBI which brought a number of stakeholders together to discuss developments in the field. This has resulted in the creation of a perspectives review and a toxicology ontology roadmap to provide guidance to the field (Hardy et al 2011a, b). As OpenTox creates a semantic web for toxicology, it should be an ideal framework for incorporating such future ontology developments, thus supporting both a mechanistic framework for toxicology in addition to best practices in statistical analysis, a combination that the field will surely profit from.

4.6 OpenTox Applications

OpenTox reached a major milestone in Spring 2010 with the development of its first two prototype applications ToxPedict and ToxCreate which were successfully demonstrated live at a workshop ahead of the first AXLR8 meeting in Potsdam. These applications have continued to be improved and were successfully used in the final hands-on OpenTox workshop in Munich in August 2011 where they received overall very positive evaluations by a variety of users. Online tutorials have also been created supporting user usage of the applications on an ongoing basis. International use in 2011 has been broadened and demonstrated by the coordinator who for example has worked with risk assessors at Health Canada and African scientists at SETAC Africa in the use of the applications in workshops, in addition to responding to a number of presentation invitations from industry and international conference organisers.

Additional application developments in the third year of the project have included the development of the Bioclipse-OpenTox integration, the Q-Edit reporting editor, the CheS-Mapper visualisation tool, and the Taverna workflow integration whose evaluation is reported in this report. These applications have demonstrated that a variety of GUIs may build upon the OpenTox foundation of resources to provide a diverse set of experiences to users through user-friendly interfaces and environments, significantly increasing the user community for OpenTox. The Bioclipse and Taverna integrations are particularly significant as they demonstrate the successful integration of OpenTox with two other major platforms, thus accelerating the capabilities and value creation of OpenTox. These developments all build and take advantage of the significant investment in the design of the framework and its open standards carried out during the research project.

4.7 OpenTox Virtual Machine

In 2011 an OpenTox Virtual Machine was developed. This allows the user to download and install a virtual machine on their computer where OpenTox applications can be run in a local environment. Hence although OpenTox supports a distributed set of resources, it may also be collapsed to run as a local stand-alone application. We brought the virtual machine software to SETAC Africa and successfully installed in on ca. 40

laptops of workshop participants. Commercial organisations have also expressed interest and have carried out initial downloading and testing because they are interested in usage behind their firewall. This development is too recent to provide concrete conclusions on its merit at this stage, but it points to an interesting direction for future development.

4.8 REACH models

A number of (Q)SAR models were built with OpenTox for a number of REACH endpoints. (See Appendix C for a representative list of models). Several of these models have been made available with validation reports through the ToxPredict application. We did not have the time and resources at the end of the project to carry out an extensive evaluation and comparison of algorithms and parameters on different datasets and endpoints. However the framework is in place to do so as a worthwhile future exercise. We propose a "challenge" should be organised in which the community could develop different models using OpenTox services applied to different REACH endpoints. The results could be reviewed and discussed, for example, at the OpenTox 2012 meeting. Such an exercise could include:

a) execution of workflows across multiple algorithms to determine best models;

b) execution of workflows involving parameter variation to determine best models;

c) evaluation of overfitting concerns;

d) optimisation of models for specific purposes, e.g., minimising false negatives;

e) determination of heuristics that could support automatic selection of best practice approaches in model building;

f) provision of detailed information on confidence and uncertainty;

g) consideration of performance of consensus models;

h) Application of (Q)SARs to mechanistic-based categories and datasets;

i) Combination of (Q)SARs with expert interpretation, read across;

j) Application of (Q)SARs with a Weight of Evidence approach.

4.9 Issues and Limitations

At the end of every research project there are always issues that the work has surfaced which provide input into future work. Here we summarise some of the main issues which we currently have discovered or recognised:

1. OpenTox is a brand new framework making use of concepts such as semantic web, ontology, REST web services, linked resources and a complete new set of APIs developed on the project, which did not exist before its start. Hence developers need to learn about the new approaches before being able to implement. Whereas we have interacted significantly with external developers and received considerable interest, there is not as yet a significant implementation by other developers;

2. It is difficult to maintain the performance of a distributed set of resources. Processes will need to be introduced to ensure the ongoing evaluation and stability of core distributed services which users and other parties rely on for their work activities. In some cases a centralisation of critical resources may be a useful approach;

3. In some cases, e.g., organisations working with in-house confidential data, users may prefer a local application, which is an area requiring future development by application developers.

4. Additional resources are needed to ensure the quality of a resource e.g., to curate a data source and to filter and label for quality;

5. Applications requiring integration of resources from multiple partners require a significant integrated testing methodology and effort;

6. A well-defined API is insufficient on its own to support developers in their development of resources to be accessed by other applications. For example, a process with check-lists for registration of a new model with the OpenTox ontology service would be useful. Automatic testing of such a new resource would also be beneficial;

7. The field of predictive toxicology is developing and expanding rapidly and involving an interdisciplinary approach of new methods in chemistry, biology, engineering and computer science. Hence a significant effort is required to incorporate support for such a broad range of datasets, protocols and models. Even more significant is the importance of communication between different disciplines who do not understand each other very well and need to learn to do so;

8. The "21st century" mechanistic approach to predictive toxicology requires a framework and methods that place numerous demands beyond the building of a "simple" (Q)SAR for a single endpoint. Through the semantic web design of OpenTox we expect to have supported its extensibility into a knowledge infrastructure involving a layer of biological concepts, but this will also require significant additional work;

9. As discussed in the Toxicology Ontology Roadmap paper, a significant amount of work is needed to organise and harmonise the set of ontologies needed for predictive toxicology. This will take a significant coordinated effort and investment;

10. Further application development, communication skills and learning methodologies are needed to provide easy-to-use fit-for-purpose tools for end users not specialised in computer science techniques.

4.10 Recommendations

We provide here recommendations for the future development and sustainability of OpenTox and its benefit and impact:

1) OpenTox should be developed further to support creation of a global interoperable infrastructure for Predictive Toxicology to aid achievement of the goals of the 21st century vision. This could be supported through a dedicated EC infrastructure project supporting the goals of a number of programs including European activities under REACH, FP7 Health, FP7 Environment, EC JRC, and IMI and international activities such as the US EPA ToxCast program and Tox21. The unique benefit of OpenTox lies in its core commitment to open standards which can support in a neutral way the goals of all stakeholders;

2) A program should be put in place to implement the Toxicology Ontology roadmap. OpenTox can play a leadership role in collaboration with other stakeholders;

3) In consultation with industry, CEFIC and ECHA, an analysis should be carried out promptly to identify steps that could be taken to optimise OpenTox for supporting the activities of industry in meeting the 2013 REACH deadlines;

4) A business ecosystem approach providing value to industry and SMEs should be developed for OpenTox in support of its economic exploitation;

5) Consideration should be given to the formation of a legal structure for OpenTox e.g., as an international organisation or foundation;

6) The potential of OpenTox to be exploited in the development of safer medicines should be pursued;

7) OpenTox has evolved a community of interested parties beyond the initial FP7 partners. This should be developed further as an active and vibrant network;

8) OpenTox partners committed at their final project meeting to a combined interest and resolve to continue to develop OpenTox beyond the end of the project. Doing so will be rewarding!;

5. Conclusions

A brief review of OpenTox history provides context to where we have currently reached. OpenTox partners held their first meeting in Basel in September 2008. At that point we had a description of work but otherwise a blank canvas for creating something new for the field of predictive toxicology that would help solve identified issues and develop a new capacity for the field. In subsequent months many discussions and explorations produced a blueprint and design for what OpenTox could be. We developed OpenTox from the ground up, nurtured a vision and philosophy, adapted or developed standards and interfaces to create what the OpenTox Framework would be. The partners then went about initial implementation. In parallel we engaged in many interactions with other stakeholders in the field that refined the vision. The outcome was an extensible semantic web vision for the field of toxicology that would provide a framework for data management, modeling and a set of linked interoperable resources. In 2009 we defined and subsequently improved a critical specification as the OpenTox API. We identified and elaborated use cases that could be implemented against. By spring 2010 we created and successfully demonstrated initial prototypes. We developed approaches and a growing consensus in the community of the importance of ontology development. In the last year of the project further applications were developed including initiatives by other developers to expand and implement the vision into their applications and platforms.

The above history set the stage for the final project workshop and conference in Munich. The conference in Munich attracted a diverse and knowledgeable audience including many principle investigators in the field, working in different areas. Hence, the group of ca. 80 scientists was strongly cross-disciplinary, cross-sector and international. With this backdrop we ran a final workshop in which 9 different OpenTox exercises and applications were run by participants in groups in a hands-on format. It is difficult to communicate the intensity of the day as many participants engaged in a variety of activities all of which had been made possible by the project. This report summarises those experiences extended through additional experiences that went before and were continued afterwards through the provisions of online tutorials.

The evaluation of OpenTox by these users provided a clear response that the project has produced very valuable and concrete results. Moreover, although still at the prototyping stage, many users were able to carry out activities with applications of significant scientific interest to them, including activities that had not been previously possible for them. The majority of evaluation answers to our questions produced answers in the excellent and very good categories, providing a sound endorsement by the user community that we have created something really valuable and worthwhile from the project. OpenTox should therefore be seen as a special success story the European Commission has supported through its research program.

A reader might ask though what is the impact of OpenTox on Europe and the world going to be? We live in a world which is under increasing pressure from its human inhabitants. We need to evaluate the risk of many chemicals in the environment, the safety of food and drugs we develop and use, the threat to climate and biodiversity of development and global warming. OpenTox is open and fundamentally a scientific collaboration and community effort. It also is a framework providing a foundation on which many applications can be developed which can have impact on the above mentioned diverse areas. It therefore provides a springboard for both academic research and industrial development to develop scientific and economic benefits for our society through supporting the creation of the safer production and use of products. The evaluation of OpenTox reported here points out that we have completed a first phase of a journey, which offers many promising roads going forward, some of which are clear and some which may provide benefits whose route is not yet clear or even expected at this point. But we have created the capability for that to happen.

6. References

(Hardy, Douglas et al, 2010) Barry Hardy, Nicki Douglas, Christoph Helma, Micha Rautenberg, Nina Jeliazkova, Vedrin Jeliazkov, 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 Gloriozova, Sergey Novikov, Natalia Skvortsova, Dmitry Druzhilovsky, Sunil Chawla, Indira Ghosh, Surajit Ray, Hitesh Patel and Sylvia Escher. Collaborative Development of Predictive Toxicology Applications, *Journal of Cheminformatics* 2010, 2:7 (31 August 2010).

(Hardy et al. 2011a). Barry Hardy (Douglas Connect and OpenTox), Gordana Apic (Cambridge Cell Networks), Philip Carthew (Unilever), Dominic Clark (EMBL-EBI), David Cook (AstraZeneca), Ian Dix (AstraZeneca & Pistoia Alliance), Sylvia Escher (Fraunhofer Institute for Toxicology & Experimental Medicine), Janna Hastings (EMBL-EBI), David J. Heard (Novartis), Nina Jeliazkova (Ideaconsult), Philip Judson (Lhasa Ltd.), Sherri Matis-Mitchell (AstraZeneca), Dragana Mitic (Cambridge Cell Networks), Glenn Myatt (Leadscope), Imran Shah (US EPA), Ola Spjuth (University of Uppsala), Olga Tcheremenskaia (Istituto Superiore di Sanità), Luca Toldo (Merck KGaA), David Watson (Lhasa Ltd.), Andrew White (Unilever), Chihae Yang (Altamira), Toxicology Ontology Perspectives, submitted to Alttex, 31 October 2011.

(Hardy et al. 2011b). Barry Hardy (Douglas Connect and OpenTox), Gordana Apic (Cambridge Cell Networks), Philip Carthew (Unilever), Dominic Clark (EMBL-EBI), David Cook (AstraZeneca), Ian Dix (AstraZeneca & Pistoia Alliance), Sylvia Escher (Fraunhofer Institute for Toxicology & Experimental Medicine), Janna Hastings (EMBL-EBI), David J. Heard (Novartis), Nina Jeliazkova (Ideaconsult), Philip Judson (Lhasa Ltd.), Sherri Matis-Mitchell (AstraZeneca), Dragana Mitic (Cambridge Cell Networks), Glenn Myatt (Leadscope), Imran Shah (US EPA), Ola Spjuth (University of Uppsala), Olga Tcheremenskaia (Istituto Superiore di Sanità), Luca Toldo (Merck KGaA), David Watson (Lhasa Ltd.), Andrew White (Unilever), Chihae Yang (Altamira), A Toxicology Ontology Roadmap, submitted to Toxicological Science, 31 August 2011, invited resubmission with minor formatting changes, 31 October 2011.

(Jeliazkova, N and Jeliazkov, V. 2011) Nina Jeliazkova, Vedrin Jeliazkov, "AMBIT RESTful web services: an implementation of the OpenTox application programming interface", Journal of Cheminformatics 2011, 3:18, doi:10.1186/1758-2946-3-18, 2011.

(Tcheremenskaia et al. 2011) Olga Tcheremenskaia, Romualdo Benigni, Ivelina Nikolova, Nina Jeliazkova, Sylvia E. Escher, Helvi Grimm, Thomas Baier, Vladimir Poroikov, Alexey Lagunin, Micha Rautenberg and Barry Hardy. OpenTox Predictive Toxicology Framework: toxicological ontology and semantic media wiki-based OpenToxipedia. In proceedings of Bio-Ontologies 2011: Semantic Applications in Life Sciences, held in conjunction with ISMB 2011, Vienna, July 2011. Reviewed and selected as invited paper in preparation for Journal of Biomedical Semantics.

Appendix A Internal Testing Results

Appendix A reports on a representative set of internal testing activities carried out on OpenTox services and applications during the final stages of the project.

A.1 ToxPredict

A.1.1 Review of ToxPredict (new GUI) by David Gallagher, 17 July 2011

Web browser: Firefox 5.0, and Google Chrome 12.0.742.112
OS: Windows XP, SP3

Start page   http://apps.ideaconsult.net:8080/ToxPredict

This a brief review of the "ToxPredict.org" application GUI and usability from the perspective of first impressions of a new (non-QSAR-expert) scientist who might want to use ToxPredict to create data for REACH purposes. It is recommended that the key issues identified here are addressed before the end-user testing and evaluation workshops in Munich.

Part 1, Summary

The Good:

- Clean, pleasing graphics and text
- Fast interactivity
- Calculations appear to complete quickly
- A good range of end-points

The Gaps:

- No apparent explanation of what "ToxPredict" is, or what it does, or who should use it.
- No mention of REACH and how ToxPredict is relevant to REACH
- Results are too cryptic with no apparent validation information, no explanations of end-points, how they are calculated, or what the results mean.
- Start page layout is not optimized for easy and intuitive use
- No apparent links or references to OpenTox
- No legal disclaimer or warning notice
- Some calculations fail with only cryptic messages

Part 2, Detailed report

1. First impressions of the GUI:

Pros: Clean, pleasing graphics and text (improvement over first ToxPredict GUI)

Cons: a. No explanation of what "ToxPredict" is, or what it does, or who should use it. Suggest adding a short tag line below the "ToxPredict logo such as "toxicity prediction for REACH, without animal testing", and also succinct sentences describing (i) what it does and (ii) who should use it.

b. No link or reference to the OpenTox project, currently ToxPredict is stated to be "Developed by Ideaconsult Ltd. 2011"

c. Drawing editor takes up too much space for a tool that is likely to be less used than the other data entry methods and so, could be hidden behind a tab (for example).

d. Page content is too wide to fit on a net book or mobile browser (partly caused by the drawing editor which could be hidden). As the browser page width is reduced, all the data entry methods except the 2D drawing tool are hidden.
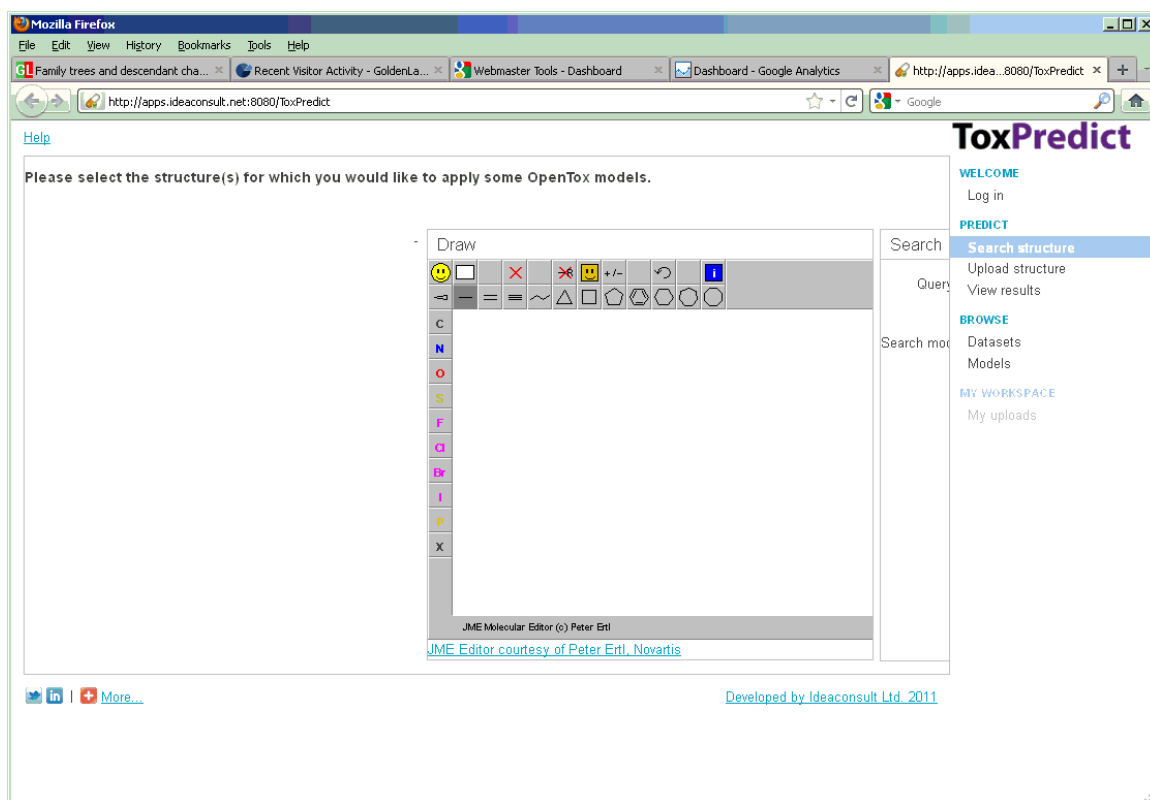
e. The "Help" menu link opens a box that is sometimes partly hidden under the Drawing editor. However, if the cursor is moved the help box disappears and the user needs to click the Help menu again to bring it back. A pop-up window might be more useful.

f. In English, pages are read from the top-left to the bottom-right, so pages normally have titles and menus set towards the top and/or left of a page where the eye is naturally drawn to. This page is unusually laid-out in that the title and menus are placed on the right side, contrary to normal interface guidelines. Suggest moving titles and menus to top left of page.

g. For both our protection and for the user's protection we should have a disclaimer and warning on the start page and results page explaining the potential limitations.

h. Some sort of progress bar or time estimate is required once the calculation is started. Currently, the users have no way to know if they need to wait seconds or hours or days.

2. First test molecule, "Trinitrotoluene"

 "Trinitrotoluene" was typed into the text box and the "Search" button was clicked, then the display refreshed with no apparent changes.

Pros: a. A "not found" message appears at the bottom right corner of the screen.

Cons: a. The "not found" message at the bottom right corner of the screen disappears after a few seconds. The location of the message may be too far away from the data entry (where the users attention is) to be noticed.

b. It is not clear why a common molecule such as Trinitrotoluene is "Not found", nor what the user should do about it. Suggest the "not found" message appears by the data entry box and remains until cancelled or clicked on. Also, some guidelines on what to try next would be helpful.

3. Second test molecule, "toluene"

"Toluene" was typed into the text box and the "Search" button was clicked, then the display changed to a second page showing the structure of toluene and a list of properties and end-point predictions.

At this point the "Run All" button was clicked

Pros: The "Run all" button is a big improvement on the having to select each calculation individually.

Cons: a. It seemed that no reference information was easily available on any of the end-points (such as links to appropriate publications and web sites)

b. A number of calculations apparently failed and cryptic error messages were encountered. Ideally, more explicit messages with information on what to do about the problem are required.

[500] Connection refused

[500] The server encountered an unexpected condition which prevented it from fulfilling the request

Access forbidden – are you logged in?

c. When the results page was resized, the results disappeared and reverted to the previous page.

4. Presentation of results

Very brief and cryptic results are presented under or alongside each endpoint. Currently, the information seems too cryptic to be useful to most people and particularly for REACH purposes. There are no explanations of end-points, how they are calculated, or what the results mean. Most importantly, there is no apparent validation information. Suggest we link in the available validation services to appear automatically alongside each result, either in full or as explicit links.

A.1.2 Review of ToxPredict (new GUI) by Roman Affentranger, 27 July 2011

Firefox 3.6.18, Windows 7 Home Premium SP1

General issues

– "Search Structure" page: "Auto detect" is described in terms of entry of the chemical. However, it is not clear what kind of search it is doing. I have the impression it's doing an exact structure search, but this is not clear. If it is doing an exact structure search, why having another "Exact structure" search type?

– "Search Structure" page: can a similarity or a substructure search be done starting with a CAS number?

– "View results" page: Some models seem to give a unit in their name, e.g., the ECOSAR LC50 fish. What's "F96-LogLC50mmol/L"? Is it the logarithm of the LC50 measured in mmol/L (equiv. to mM, assuming L stands for liter)?

– What's the meaning of the color of the little activity-indicating square in the top-right corner changing from white over yellow to red (and then often staying red for a long time)?

– "View results" page: Explanations for what the models are about, what they are predicting, is needed. E.g., what's the "QSAR SRC KOWWIN fingerprints AD". All one gets when mousing over is the model URI (http://apps.ideaconsult.net:8080/ambit2/model/130825). The model URI is not clickable, but needs to be copy-pasted to the address bar to go to the model URI.

– "View results" page: What are the units of the model predictions? E.g., What's the unit of the Caco-2 prediction result? It's quite tedious having to read the full paper to figure that out (and still, one doesn't know for sure that the model is using the same unit).

– "View results" page: The verbose explanations for the ToxTree models, START and SmartCYP make the page very long and should be hidden by default.

– "Log in" / "View results" page: When not logged in, I get the nice message "Access forbidden – are you logged in?" for some of the models. When logging in, I think the information where to get an account should be more prominent. It's easily missed.

– "View results" page: For IST/lazar models, the confidence is reported, but it is not explained what it is.

– "View results" page: SMARTCyp: No information what the atoms refer to (e.g., SMARTCyp.Rank1.sites C6,C14: which atoms are these?)

– "Log in": After logging in, one lands at the "My account" page, and not where one has been before. If you didn't pay attention before that the starting page when entering "toxpredict.org" is the "Search structure" page, or that after having entered a structure, one is at the "View results" page, one might not know how to get back to where one was before.

– "View results" page: Even when logged in, I get "Access forbidden – are you logged in?" for both "OpenTox model created with TUM's kNNregression model learning web service" models (http://opentox.informatik.tu-muenchen.de:8080/OpenTox-dev/model/TUMOpenToxModel_kNN_114 and http://opentox.informatik.tu-muenchen.de:8080/OpenTox-dev/model/TUMOpenToxModel_kNN_92). When going to "BROWSE" -> models and trying to "show" this model, I don't get more than:
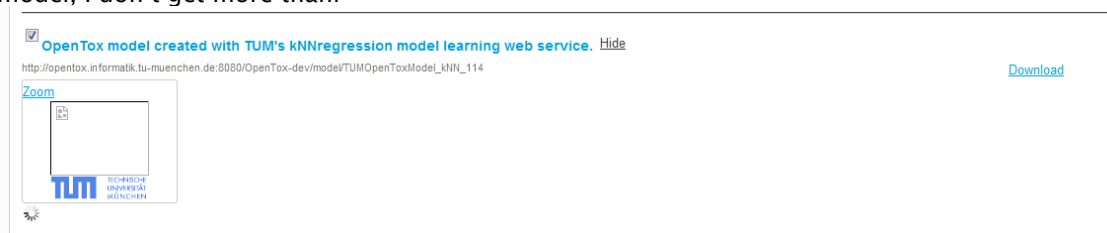
at this point, nothing more happens.

UPDATE 2011-07-28: Today, these models appear to be running, but are not producing results (nor errors).

- "View results" page: http://opentox.ntua.gr:8080/model/2bf22fab-a9bf-4295-8172-1d18d4476da0 and http://opentox.ntua.gr:8080/model/6a3dff28-b3f3-426d-9428-b7c31d9eb599 give "Server returned HTTP response code: 500. Only http://opentox.ntua.gr:8080/model/536ce048-b78b-4fb3-8f0f-1c85447bfb41 is working OK.

- The error messages of the two NTUA models given above don't have line breaks, and can thus not be read fully:

- The MLR model for Exp LogKow takes a long time (20 minutes +) without producing a result (or an error).

- The model http://opentox.ntua.gr:8080/model/536ce048-b78b-4fb3-8f0f-1c85447bfb41 gives some output, but it's unclear what these mean, and also what the model actually is about. Only from the "BROWSE -> Models" page it seems it's a Caco-2 model…

- When moving away from the "View results" page to e.g. the "BROWSE -> Models" page to get some information about a model, everything's lost when returning to "View results". That's a problem even if the results of previous predictions are stored, because there were some models that were still running. When returning, there's no indication that they're still running. So, if there's a model that looks like it's running for 20 minutes, I better not move away from the "View results" page.

- "Browse Datasets" page: When browsing a dataset, four molecules are shown, the rest is paged. The four molecules cannot be shown entirely, so there's a vertical and horizontal scrollbar:

That's OK, since I have a rather small window (well, full screen on a 1366x768 laptop screen). However, even if I move the window to a larger 1680x1050 screen the scrollbars remain:

Strangely, when moving to the next page, the scrollbars are gone, even on the smaller laptop screen. Even more strangely, when moving back to page 0 (where I had the scrollbars just before), they're gone as well.

- After having clicked "Browse Dataset" in the "Datasets" tab on the "View results" page, I want to go back there to check the next dataset. The browser's "Back" button, however, brings me to the "Predictions" tab of the "View results" page. So, if I want to browse a number of datasets in which "my" compound can be found, I have to do a lot of clicks: "Datasets" tab -> "Show" for a given letter or "Show All" -> "Browse Datasets" -> "Back" -> "Datasets" -> "Show" -> "Browse Datasets" etc.

- "Browse dataset": It's not very clear what the checkboxes "Show Identifiers", "Show Data", and "Show Predictions" are for. For the TCAMS dataset (http://apps.ideaconsult.net:8080/ambit2/dataset/584486) nothing happens when I tick of "Show Identifiers" or "Show Data". Also, when clicking on "Show predictions" I initially thought nothing happens (I didn't see the little square appearing at the top-right).

- Ticking off "Show Predictions" takes a very long time for something to happen for the TCAMS dataset. UPDATE 2011-07-27: Yesterday I clicked "Run models" for this dataset. Today it takes even longer to "Show Predictions". It's been running for almost one hour without anything happening.

- "Browse dataset": It's not entirely clear what the "Run models" button does. Does it run all models on all compounds in the dataset? Or just the ones on the current page?

- "Browse dataset": What's the purpose of the "Select Structures" line? I think I remember having been able to click there to select/unselect individual molecules. However, this is not working anymore. Why is it there? How can I select/unselect compounds in a dataset?

- "Browse dataset": Clicking "Run models" informs me that the process started and gives the URI of a dataset (that will be created?) – however, the message disappears before I get the chance to e.g. note the dataset URI. So why not leave it away entirely? NICE: the progress bar showing %-completion!

- "Browse dataset": If I run all models on a dataset and open ToxPredict on another browser tab, query the same molecule, go the "Browse dataset" of the same dataset, I get the progress bar, but without the "XY % completion" indication.

– "Browse dataset": When clicking on a molecule drawing, the following appears:

But the conformer URI cannot be copied or clicked. As soon as the mouse is moved, it disappears (although the mouse pointer changes from arrow to a hand when moving over it quickly. It's not clickable, though).

– "Browse dataset": When downloading the dataset (tested for the TCAMS dataset: http://apps.ideaconsult.net:8080/ambit2/dataset/584486) I get a very weird CSV (and SDF) file (haven't tested the other formats). In terms of properties/features, all I get is the compound URI, the MW, Tanimoto and AppDomain Tanimoto, caco2, some micronucleus things (including details), Exp LogKow by MLR model, and three more prediction features. That is, most of the prediction features are missing, as well as the actual contents of the dataset. Also, when opening the CSV in Excel, the columns are not OK. Also, I only get 10000 compounds (and not all the 13500 compounds of the dataset).

– "Browse dataset": The download doesn't let you choose if you want to only get the dataset with its inherent features, or also all the prediction features.

– "View results" page, "Predictions" tab: After having run all predictions, it's not quite clear what to do to save the results. The "Download" link beneath the structure diagram has been there from the beginning (and doesn't really give proper files, above). There's no "Save results" or "Download results", or a checkbox what to download (data in dataset, predictions, etc.).

– "BROWSE" -> "Models": To get information about a model, several "Show" links need to be clicked, each associated with quite a delay. First, the model needs to be expanded by clicking "Show":

Clicking on "Show" for "Y (Predicted) opens a box with "Column 1" and again an OpenTox algorithm for which one needs to click "Show" (same for endpoint).

– "BROWSE" -> "Models" (probably also applicable to other pages): When scrolling down on the page such that the top is not visible anymore, the red-box indicator of activity is not visible anymore. That is, it's displayed in the absolute top-right corner of the screen instead of the top-right corner of the visible part of the page. Thus, when clicking something that doesn't have an immediate effect, it's not clear if just nothing's there, or if something's still in progress.

Compound-specific issues

1. Using OC(C1CN(CCc2ccccc2)CCC1(O)c1ccccc1)c1ccccc1 and "Auto detect"

- "View results" page, "Predictions" tab: I get a value of -4.59 for the ECOSAR fish toxicity. If my interpretation of the unit is correct (see comments on units before), this would mean that the LC50 is 10-4.59 mM, that is, 10-7.49 M, that is, 25 nM? In comparison, with the IST fish toxicity model I get 0.0103 mmol (again, the mmol is deduced from the name of the model. Tis unit doesn't make much sense, though. Is it mmol/l, or mmol/ml?). Assuming the "mmol" in the title of the IST fish toxicity means mM, that's a difference of 400 to the ECOSAR model.

- "View results" page, "Datasets" tab: I get two datasets: "T(1)" and "_(58)". These are weird names of calling them. Why do I have to click on "Show" next to "T(1)" to see the name of the dataset (TCAMS Malaria Box)? With only such few results, it's weird having them grouped this way.

- "View results" page, "Datasets" tab: Update 2011-07-28: Expanding the 58 datasets under "_(58)" gives the following Firefox message:
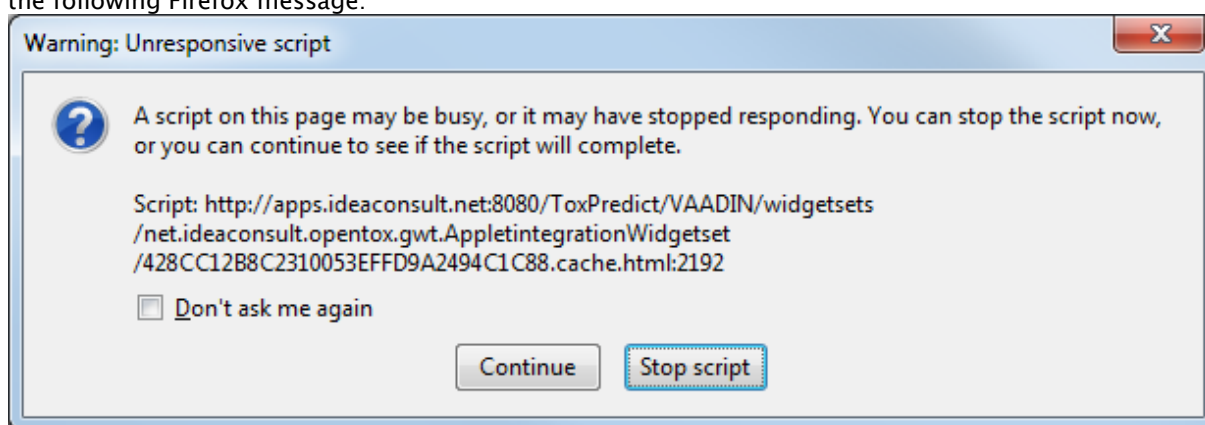
and the red square in the top-right corner indicating activity keeps being there.

- "View results" page, "Datasets" tab: Apparently, under "_(58) there are 58 datasets. They are all called "_ambit2_053567f4-2de5-4d0a-8b25-3fff50e433a521285" (the string after _ambit2_ varies). What are those?

- "View results" page, "Datasets" tab: Clicking on "show" next to "_ambit2_053567f4-2de5-4d0a-8b25-3fff50e433a521285" (one of the datasets inside the "_(58)" is taking a long time for anything to happen (several minutes). Clicking on the "stop"-cross of firefox pops up the red "communication problem" box.

- "View results" page, "Datasets" tab: Clicking on "show" next to "_ambit2_053567f4-2de5-4d0a-8b25-3fff50e433a521285" now seems to be doing something. An empty white box appears, and Firefox tells me that it is still reading apps.ideaconsult.net (see screenshot below). Clicking on "Show" for another one of the "_" datasets, and the same thing happened (empty box). Clicking on "Show" next to the TCAMS Malaria Box dataset works, and I get a box with properties.

- When going to "Browse Dataset" in the "View results" page, "Datasets" tab for the TCAMS Malaria Box, in the page that opens "my" molecule is not there, nor do I know where in this dataset "my" molecule is.

- "View results" page, "Datasets" tab: trying to "Browse dataset" for one of the "_" datasets doesn't work. Initially, I had the impression nothing's happening. Then I saw the little red rotating square at the top-right of the screen, indicating that something's going on. That was difficult to spot. Also, nothing ever happened.

- "View results" page: When clicking "Download" beneath the structure diagram, the same problems apply as described previously. In general, it is not clear what this download does.

- The "Browse all" link is confusing if there's only one compound found.

- "BROWSE" -> "Datasets" yields the same page as clicking on "Browse all" beneath the structure diagram on the "View results" page:

In particular, there's no data shown, no features, identifiers, etc. Even ticking off "Show identifiers" and "Show data" didn't change anything.

2. Using OC(C1CN(CCc2ccccc2)C1(O)c1ccccc1)c1ccccc1 and "Auto detect"

It's the same smiles as in the previous example, just with two carbons removed in the middle (without thinking about whether that makes sense or not).

- Keeping "Auto detect" for this SMILES and clicking "Search", I get the little square indicating activity, changing color from white over yellow to red. Then the screen flickers, and that's it. No error message, no "NOT FOUND" message, nothing. Same if I choose similarity search with tanimoto=0.9.

- With similarity and tanimoto=0.6 I get four hits. Clicking on "Browse all" I can see them, but can't do anything more:

For example, I can't deselect any, or show properties. To see the data associated with these I have to select each molecule individually, go to the "Datasets" tab in "View results", expand the datasets and browse them.

Drawing a structure

I was drawing a structure similar to the one used in the previous example, but slightly different.
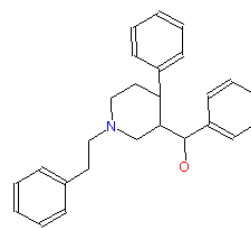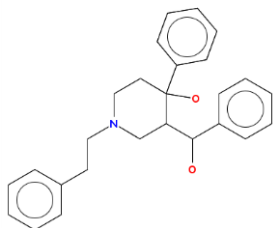
Structure in previous example: Structure I drew:
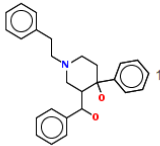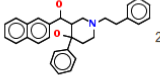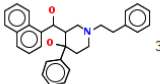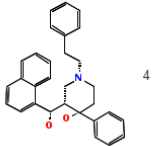
With a Tanimoto of 0.9 I get four hits:
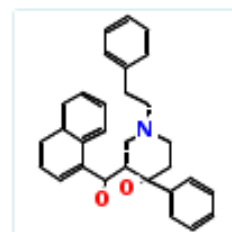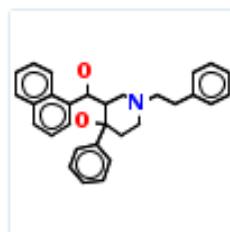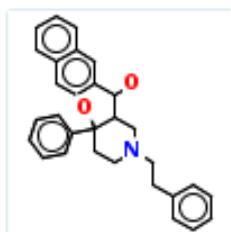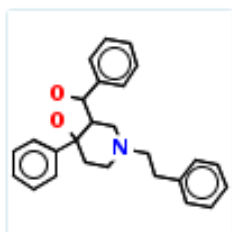
- There is no indication if I found exactly the structure I was drawing, or just similar ones (I tried drawing the structure used before, and although it found it, there was no indication that it was exactly the structure I drew).

- I couldn't see the similarity to my drawn structure of the found hits. Clicking on "Browse all" beneath the structure diagram on the "View results" tab gave:
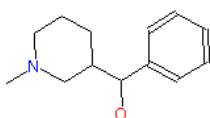
but the "Similarity" row was empty.
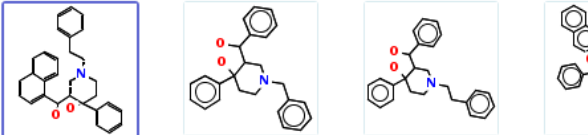
Substructure Search
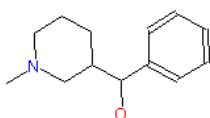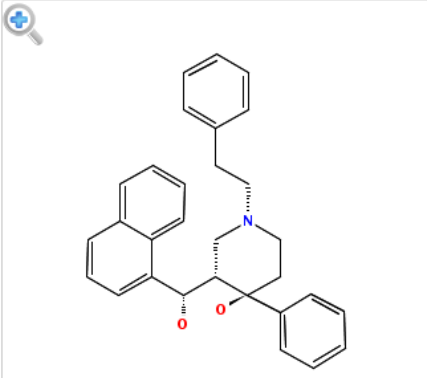
I drew   and performed a substructure search.

  - I got a number of hits, but it didn't say how many.

  - The "Browse all" link beneath the structure diagram had disappeared:

  - UPDATE 2011-07-27: "View Results" tab: When I click on "Download" beneath the structure diagram, a new tab opens, says "Loading…", but nothing ever happens (both in IE – not logged in – and on Firefox, logged in).

3. Using Cl.NC(=N)Nc1ccc(Cc2ccc(NC(N)=N)cc2)cc1

- – "View results" page: when running all models, the list gets very long, and the SMARTCyp results cannot be seen anymore (and one cannot scroll down):

A.1.3 Review of ToxPredict (new GUI) by Roman Affentranger, 23 August 2011

Firefox 3.6.18, Windows 7 Home Premium SP1

Unless otherwise mentioned, I will be working on this molecule:
(C1=CC=CC=C1CCN4CCC(O)(C2=CC=CC=C2)C(C(O)C3=CC=CC=C3)C4)
Dataset URI: http://apps.ideaconsult.net:8080/ambit2/dataset/647254
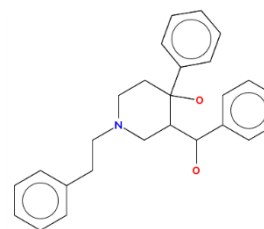
Starting Page toxpredict.org

- **NICE**: when drawing a structure, one gets the SMILES in the Query field. I really like this feature.
- The instruction "Please select the structure(s) for which you would like to apply some OpenTox models" sounds like one can select structures. I would rather call it enter.
- As far as I can tell, "Exact structure" search is a subset of "Auto detect". That is, "Auto detect" is the same as "Exact structure", except that it allows for some additional text queries. Couldn't these two search types be merged?
- **BUG**:   Mousing over the IdeaConsult logo, it says IdeaConsult was in **Buligaria** ;-)
- The vertical space around the ToxPredict logo at the top is a bit tight.
- As discussed before, the "Help" link should be moved away from the top-right corner (to the menu?), as it is hard to see inside the purple gradient.
- **NICE**: I like how one gets mouse-over information for the OpenTox logo and for the IdeaConsult logo.
- **BUG**: Mousing over the menu items, I get "List of uploaded chemical structures and datasets. Enabled only for registered users." for every menu item except for "Log in" and "My Workspace".
- **NICE**: although one has to click "Search structure" again after logging in (as mentioned in the last report), the structure one had input before logging in is still there when returning to the "Search structure" page.
- The "NOT FOUND" message box at the bottom-right of the screen is too small, I think. Could this be made to appear more close to the "Search" button? It's where I'd expect such an answer.

Help information: "In the first step of your ToxPredict workflow you are invited to select the structure(s) for which you would like to apply some OpenTox model(s). The **Search**, **Draw** and **Upload** tabs below provide different ways for structure selection: "

- Again, I find "select" confusing. What does one select from?
- The help text talks of "The Search, Draw and Upload tabs", but these are not really tabs, are they? In particular, the Upload is quite different.
- The help information does not say that it is necessary to be logged in to upload a structure

Upload structure page

- **NICE**: the Upload button changes depending on the choice of the radio button.
- The Upload page looks so similar to the search page that they may be merged. Adding the two radio-button options of the Upload page to the Search page, and having the "Search" button change to "Upload" when selecting an uploading radio button is all that's needed. The uploading functions could be grayed out for non-logged-in users. I think it would make the interface easier.
- Uploading worked nicely, and I got to the "view results" page as with a search.

My uploads page

- I don't know what else is planned to be there, but the way it is now the information seems a bit redundant: At the top it says "Uploaded datasets", underneath that – in small font – it says "Datasets

uploaded via ToxPredict", underneath that it says – again in bold blue – "Drawn and uploaded via ToxPredict". So, basically it's three times the same information.

- **NICE**: it shows the date of uploading. Good to have that for reference.
- Not so sure how useful it will be when having uploaded many things. Since the page only gives the URI of the created dataset, I have no clue which URI corresponds to which molecule/dataset without inspecting it. Maybe there should be an option to give a title to an upload, and this title would then be displayed in the "My uploads" page…
- One can check "Dataset access rights", receiving "No policies found". However, I couldn't see an option to define a policy when uploading (like it is possible in Ambit).
- The "Predictions" link on the "My uploads" page brings me to the "View results" page. That's a bit confusing, having different names for the same thing…
- **NICE**: Still, being able to jump directly from any dataset in the "My uploads" to the "View results" page, without having to enter the dataset URI into the search box, is very nice! It's just the the meaning of the "Predictions" link is not so clear. To me it sounds more like the "show predictions" feature in the "Browse datasets" view.

View Results page

- **NICE**: the "Run All" button!!!
- The "MLR model for caco2" gives a 500 error.
- The output of the "OpenTox model created with TUM's J48 model learning web service for Micronucleus Data" is of limited use. All I get is a huge number of descriptors, but no prediction…
- Weird: TopoPSA (from CDK) in TUM's J48 model is 20.31, while TopoPSA is 40.54 in TUM's kNN models.
- There is still the discrepancy between the IST Fish Toxicity model and the ECOSAR fish toxicity model, predicting LC50s that differ by a factor 20. Maybe that's OK…
- I still think that the "Verbose explanations" should be collapsed by default, with a "<u>Show</u>" link, like it is present in many other places in ToxPredict.
- The "Datasets" tab shows 173 datasets for this compounds beginning with "_" (these were the predictions, right?). Clicking on "Show" doesn't yield anything there.

Browse Datasets page

- **NICE**: the "Run models" button! I really like this, being able to run a bunch of models on a whole dataset!
- When showing all predictions, I understand that it only shows the predictions that are cached. As far as I understood, only predictions created by models in Ambit are cached. It is a bit confusing to the user, though, not knowing this. If a user runs all models in the "View results" page, then goes to "Browse datasets" and ticks of "Show predictions", how can the user understand why the predictions for some of the models are not shown?
- When running models, are they run serially? That is, if one model takes a long time, all the others are blocked, as well? If so, that's a bit unfortunate (there's always the MLR model for ExpLogKow taking forever…

Browse Models page

- The dropdown list of Endpoints at the top should not be semi-transparent. It doesn't help being able to see what's below, but it does make it less clear what's in the dropdown.

- The "Select/Unselect" tickbox takes quite some time to react (about 3–4 seconds). The first time I used it I thought it didn't work, and clicked a second time. Then I got really confused, because by the time I clicked again, the system had reacted to my first click, so I thought it was doing the opposite of what I'd expect (ticking **on** "Select/Unselect" ticks **off** all models, while ticking **off** "Select/Unselect" ticks **off** all models). This was confusing. Only later I realized that it just takes a couple of seconds to react. The same problem happened with one participant at the workshop in Munich.
- **NICE**: I think it is a nice feature of ToxPredict, being able to work with just a subset of models throughout. However, the feature is not easy to find (and to understand). E.g., the menu item is called "Browse models". From this one doesn't expect being able to also select/deselect models with a "global" effect for ToxPredict. Even when on the Browse Models page, when one sees that there is an option to select/deselect, it is not clear what the effect of selecting/deselecting is.
- **BUG**: When unselecting all models, then moving over to "View results" (no models are shown now), then moving back to "Browse/Models", the "Select/Unselect" box is ticked on, although no models are selected. Clicking the box makes the check-mark disappear, but that's all that happens. I think the "Select/Unselect" box should be unchecked in such a case.
- Maybe it would be better to have, instead of a single "Select/Unselect" check box, a
  Select <u>All</u> <u>None</u> <u>Inverse</u>
  choice (or something similar, e.g. buttons for these choices, etc.).

Menu

- It's nice being greeted with a "WELCOME, <username>" when having just logged in, but it's strange having this remain there all the time at the top of the menu ;)
- As mentioned before, I'd combine "Search Structure" and "Upload structure"
- I'd rename "View results" e.g. to "Apply models", or "Predict toxicities" or something like that. Being called "View results" is strange since one has to click on "Run all" every time when returning there.

General

- The "Help" link seems to be context dependent, ideally displaying different things on different pages. However, it shows nothing on most pages.
- It should be mentioned that the help is context dependent. Users might land at the ToxPredict starting page, read the help, start moving around and get lost without ever checking "Help" again because they think they've already read it.

A.2 ToxCreate

A.2.1 Review of ToxCreate by David Gallagher, 19 July 2011

Web browser: Firefox 5.0, and Google Chrome 12.0.742.112
OS: Windows XP, SP3

Start page http://toxcreate.org/

This a brief review of the "ToxCreate.com" application GUI and usability from the perspective of first impressions of a new (non-QSAR-expert) scientist who might want to use ToxCreate to create a predictive model for REACH purposes. It is recommended that the key issues identified here are addressed before Beta testing and the end-user testing and evaluation workshops in Munich.

Part 1, Summary

The Good:

- Clean, pleasing graphics and text

- Easy-to-navigate tabbed interface

- Tag line and text explains purpose of the application

- Calculation progress bar provides useful user feedback

- Calculations appear to complete reasonably quickly with the test-set used

- Detailed test results easily accessible

The Gaps:

- No mention of who should use it.

- No mention of REACH and how ToxCreate is relevant to REACH

- Apparent lack of OECD compliant validation report

- No explanation of what the results (numbers) mean

- No legal disclaimer or warning notice

- Lack of an SDF file input format

Part 2, Detailed report

1. First impressions of the GUI:

Pros: a. Clean, pleasing graphics and text

b. Easy-to-navigate tabbed interface

c. Tag line explains purpose of the application. Suggest rewording tag-line to

"Creates computational models to predict toxicity". Also, suggest moving the "ToxCreate" logo to the left of tag line or above the tag-line so they are more closely associated.

d. Text on "Create" page explains purpose, but suggest rewording it to

"This service creates and validates new "Classification" and "Regression" structure-activity models from your own experimental data, for toxicity prediction including REACH purposes. The models can be used to predict toxicity of new chemicals and so reduce the need for animal testing. The following methods are currently available: Lazar, ..."

"To create a predictive model, you will first need to upload a set of training data that includes chemical structures and their measured toxicity values, in one of the following file formats: Excel, CSV, SDF,...."

Cons: e. No explanation of who should use ToxCreate, suggest adding: "Who should use ToxCreate: scientists or technicians with at least some basic understanding of toxicity and statistical methods."

   f. For both our protection and for the user's protection we should have a disclaimer and warning on the start page and results page explaining the potential limitations, such as: "Disclaimer: ToxCreate uses state-of-the-art published and tested algorithms and methodologies with full validation information. However, just as with experimental measurements, computational predictions are subject to varying degrees of accuracy and uncertainty, so please read the full report carefully, particularly the validation information. No liability is accepted for any inaccuracy in predictions."

   g. The lack of the popular SDF input file format is a major gap.

   h. Neither the brief report on the "Inspect" page or the detailed report offer any explanation of what the various numbers mean or how to interpret them. This would restrict the use of the service to experts only (see an example in 2.b below)

   j. Currently, it seems that neither the detailed report nor the brief report meet the OECD guidelines for validation, making it of limited value for REACH, etc. Suggest adding full OECD compliant validation report with explanations where possible, and format the validation report according to the OECD sub-headings, i.e.

   1. Defined Endpoint
   2. Unambiguous Algorithm
   3. Defined Applicability Domain
   4. Goodness-of-fit, robustness and predictivity
   5. Mechanistic interpretation (if possible).

   k. Suggest adding the following items to the home page

   - ToxPredict logo and link
   - OpenTox European funding details link
   - ToxCreate software version, release number and date

2. First test set, "hamster_carcinogenicity.csv"

Pros: a. Calculations appear to complete reasonably quickly with the test-set used and without any errors.

Cons: b. When the results were completed, a long list of other results were also displayed, presumably from previous unregistered users. This can cause confusion, so suggest only the results from the current user are displayed (i.e. create a new temporary user account for each non-registered user)

   c. When the model was used to make a prediction the result was provided with a confidence level number. The confidence number needs some explanation as to what it means and the typical acceptance levels (for example: 0 - 0.5 = meaningless and unreliable, 0.5 - 0.8 reasonably reliable, > 0.8 = good reliability)

3. Second test set, "hamster_carcinogenicity-one.csv"

This test set contained only a single compound and so should be invalid for model building purposes.

The model creation failed with the following message "Dataset creation failed with private method `split' called for nil:NilClass". A more useful message for the user would be "There was only 1 chemical(s) in the training set which is insufficient to build a predictive model. A minimum of 6? training compounds with data are required"

4. Third test set, "hamster_carcinogenicity-bad-data.csv"

This test set contained randomized and invalid data and so should be invalid for model building purposes.

The model creation seemed to hang and did not complete after one hour. The progress bar (Creating prediction model) was stopped at 15% during this time. Suggest fixing this apparent bug.

5. Presentation of results

The presentation of results in both a summary format and detailed format is very helpful.

However, the apparent lack of validation information that meets the OECD guidelines is a drawback for REACH purposes. Some suggested improvements:

– Add full validation information compliant with OECD guidelines (see 1j above)

– Add links to detailed information and published papers on the available methods (e.g. Lazar)

– Provide more explanation of what the key numbers actually mean

A.3 Taverna

A.3.1 Review of OpenTox Taverna Workflows by Barry Hardy, 5 and 7 August 2011

Description of Software Tested

Initial evaluation of the Taverna workflows of Joerg Wicker.

Purpose of Test

To carry out a first impressions evaluation of Joerg's Taverna workflows to provide feedback to the developer.

Instructions:

Taverna 2.3: http://jcave.de/taverna/examples/workbench-distro-2.3-SNAPSHOT-bin.zip

To run: execute taverna-debug.bat

Initial:

Download address: http://jcave.de/taverna/examples/ where one finds three Taverna workflow files:
- handle_task.t2flow
- list_tum_alg.t2flow
- upload_and_merge_datasets.t2flow

To open them, you have to download them and open them in Taverna using File -> Open Workflow. list_tum_alg.t2flow lists all TUM algorithms. To run a workflow, click the Run button (the green triangle in the symbol bar). Taverna will switch to the results perspective. The resulting list of algorithms will be in the bottom fields, click the result on the left to see the list. upload_and_merge_datasets uploads 2 data sets and merges them to one which is downloaded in SDF format.

To create a workflow, clock on New workflow, the diagram should be empty then. You can add new REST elements by right clicking on the right area and click "Add REST service". You can configure the REST service by the configuration dialogue then. It is rather easy, like writing a cURL call.

Description of Test Results

OS: Windows 7. Taverna 2.3:

http://jcave.de/taverna/examples/workbench-distro-2.3-SNAPSHOT-bin.zip

First test (5 Aug):

For the data integration workflow I am getting XPath errors:
 "Current configuration of this XPath activity has some missing namespace mappings for the specified XPath expression. This is a severe problem and prevents from execution of the XPath activity.
 This may have happened if you have typed in (or pasted) the XPath expression, but did not manually add the required XML namespace mappings. "

However, if the XPath expression was generated automatically (by selecting desired element from the example XML tree structure),this may indicate a problem with the XPath Activity plugin itself.

Second Test (7 Aug):

Running handle_task.t2flow worked, but it was not clear what this did.

Running list_tum_alg.t2flow I got the following results:

http://opentox.informatik.tu-muenchen.de:8080/OpenTox/algorithm/kNNclassification
http://opentox.informatik.tu-muenchen.de:8080/OpenTox/algorithm/J48
http://opentox.informatik.tu-muenchen.de:8080/OpenTox/algorithm/kNNregression
http://opentox.informatik.tu-muenchen.de:8080/OpenTox/algorithm/PLSregression
http://opentox.informatik.tu-muenchen.de:8080/OpenTox/algorithm/M5P
http://opentox.informatik.tu-muenchen.de:8080/OpenTox/algorithm/GaussP
http://opentox.informatik.tu-muenchen.de:8080/OpenTox/algorithm/LR

http://opentox.informatik.tu-muenchen.de:8080/OpenTox/algorithm/BayesNet
http://opentox.informatik.tu-muenchen.de:8080/OpenTox/algorithm/IntervalEstimator
http://opentox.informatik.tu-muenchen.de:8080/OpenTox/algorithm/FTM
http://opentox.informatik.tu-muenchen.de:8080/OpenTox/algorithm/gSpan
http://opentox.informatik.tu-muenchen.de:8080/OpenTox/algorithm/FTM/{smiles}
http://opentox.informatik.tu-muenchen.de:8080/OpenTox/algorithm/gSpan/{smiles}
http://opentox.informatik.tu-muenchen.de:8080/OpenTox/algorithm/CDKPhysChem
http://opentox.informatik.tu-muenchen.de:8080/OpenTox/algorithm/JOELIB2
http://opentox.informatik.tu-muenchen.de:8080/OpenTox/algorithm/InfoGainAttributeEval
http://opentox.informatik.tu-muenchen.de:8080/OpenTox/algorithm/PrincipalComponents
http://opentox.informatik.tu-muenchen.de:8080/OpenTox/algorithm/ChiSquared

Running upload_and_merge_datasets.t2flow I obtained an error message asying that input values are not provided.

I do see this time that each task box has an interesting "Show details" where information on task and annotations are provided. It would be useful to document tasks using annotations.

Conclusions and Recommendations

More documentation, help and tutorial is needed to understand what this is doing or what information is needed or being generated or could be viewed and how to see the end results.

A user having two datasets to merge will need guidance on how to setup the datasets, how to understand and setup the workflow and its parameters, and how to access and use the results.

Third Test: LoMoGraph

Initial: please find attached a Taverna workflow example using LoMoGraph for model learning and prediction. The workflow comprises:
1. Feature calculation for train and test set
2. Model learning and validation of a LoMoGraph model
3. Application of the learned model to a test set

The workflow should be runnable as it is. The only thing you have to do is to reset the token in our web-services, since it might be expired.

First Clarification:
Some comments to the Taverna workflow that consumes LoMoGraph. I attached an image that makes clear that the workflow can be divided into 4 parts. In the black part CDK descriptors are calculated for both training and test set. In the red part LoMoGraph is validated using Martin's validation server. The task returned from Martin can be retrieved from output port "debug2". Get the task and put it into a browser. In the green part a LoMoGraph model is learned on the training set. And in the blue part, the learned model is applied on the test set. Predictions or the final dataset containing the predictions can be retrieved by querying the task returned in output port 4.

Some further comments:
- Blue boxes are web services, e.g. LoMoGraph
- Brown boxes are beanshell scripts, e.g. for trimming URIs (deleting newlines)
- Grey boxes contain strings, e.g. input parameters for web services
- Lilac boxes concatenate string
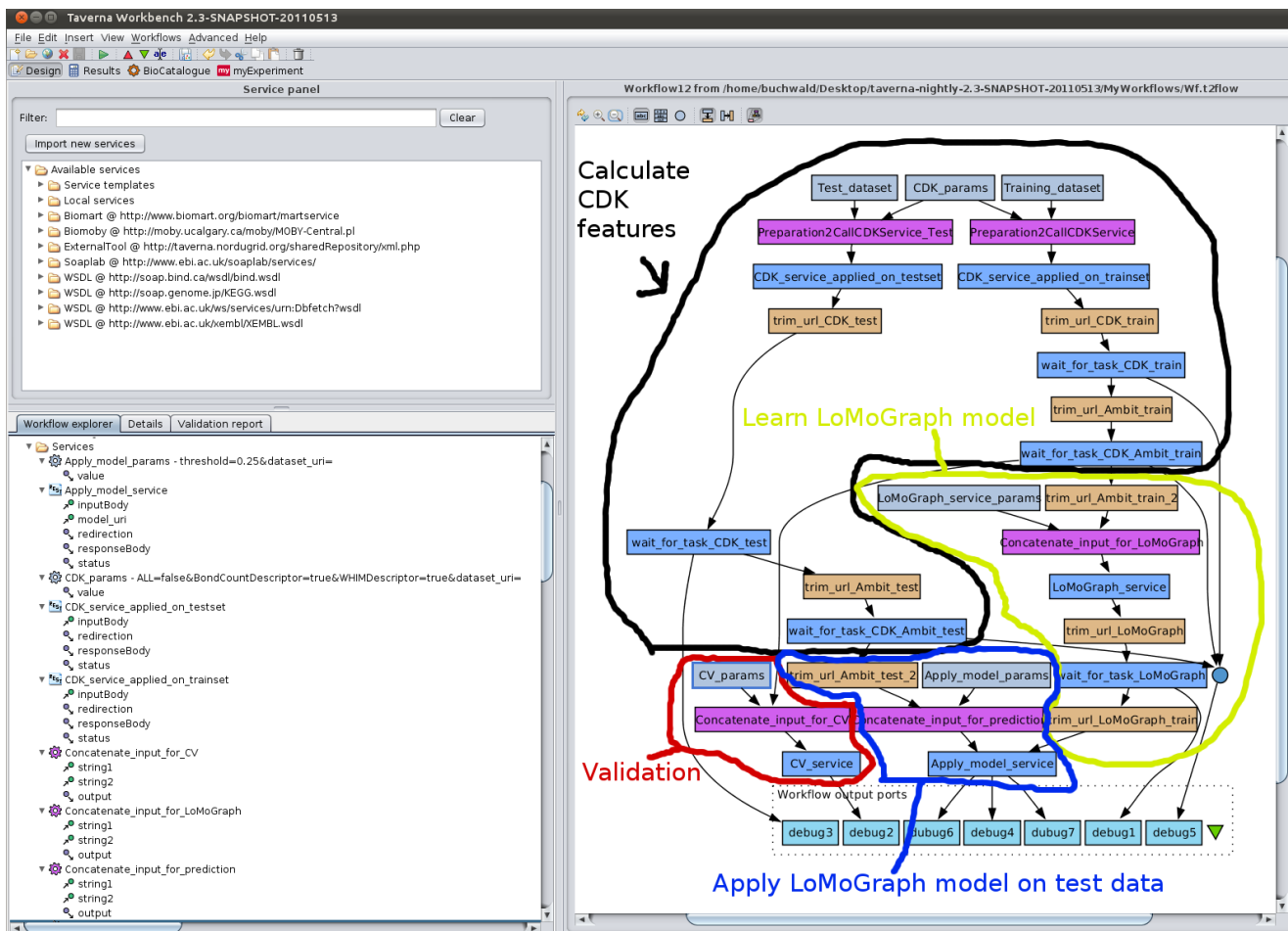
First Test (5 Aug):

I am getting status errors: e.g., Processor 'wait_for_task_LoMoGraph' – Port 'status' and not sure I am getting any results

Second Test (7 Aug):

Running: WorkflowLoMoGraph 110805.t2flow

I imported and ran the workflow. All the boxes turned grey and it "completed".

You would have to be an "expert user" on Taverna and OpenTox and have tutorials/experience to understand what this workflow is doing, let alone modify it. Apart from that though, at a first time try, one has no idea at all what this workflow does, what is the input and output, what the endpoint or model is doing, and I could not view any results. This will need detailed guidance/tutorial to have a chance of following.

Diagram provided gives one some general idea but one would need more granularity.

Conclusions and Recommendations

More documentation, help and tutorial is needed to understand what this is doing or what information is being generated or could be viewed or even what it all means.

A.3.2 Review of OpenTox Taverna Workflows by Roman Affentranger, 24 August 2011

Download:

I followed download link in instructions (http://www.taverna.org.uk/download/). I had to assume that I need the "Taverna Workbench". The comment in the instructions, that 2.3. is a development version that has to be downloaded from nightly builds, isn't very helpful without giving a hint where to get it…

NOTE: as of 24 August, version 2.3 is the current release. The instructions should be updated.

Downloaded the Installer for Windows (http://launchpad.net/taverna/t2/2.3.0/+download/taverna-workbench-2.3.0-installer.exe).

Installation:

Installation was unproblematic using the installer mentioned above. The instructions at http://opentox.informatik.tu-muenchen.de/trac/TUMOpenTox/wiki/Taverna seem incorrect for Windows systems. With the installer, one gets the application in the start menu. No need to navigate to the Taverna folder to double-click a file. The file, to be precise, also isn't called "taverna.bat", but "taverna.exe".

List all TUM algorithms

- The online instructions are incomplete. It doesn't say that the workflow should be downloaded from the "Attachments" at the bottom of the web page (nor does it say which one it is. It's rather easy to guess that it's list_tum_alg.t2flow.
- I downloaded list_tum_alg.t2flow and opened it in Taverna.
- Ran it OK

Upload and merge data

- Again, no reference to the downloadable workflow
- Opened upload_and_merge_datasets.t2flow
- Selected two .csv files as input. They have different features…
- This didn't work, but it took me a bit to figure out that it only accepts SD files (I though an SD file was just given as an example…)
- Used two SD files (one and two compounds) as example.
- This worked fine
- **QUESTION**: if it's two SD files on a local computer that are to be merged, isn't it a bit an overkill to upload both datasets to AMBIT and then merge them? The files could simply be concatenated and then uploaded…
- **QUESTION2**: cleanup of the individual files, is that done?

Handout

- The handout covers creating from scratch the "get list of algorithms" workflow. That worked fine
- The "Upload data set to Ambit" workflow is described quite OK (haven't tested it, though). However, it would be nice to know why certain settings must be made, as opposed to simply be told to apply such and such settings. E.g., is it mandatory to set the "Send data as" to "binary"? or just better? Why does one have to uncheck "Escape URL parameter values"? etc.
- The "Learn and Evaluate a Model" example is missing a download link for the workflow.
- The "Learn and Evaluate a Model" example states one might have to reset the token, but it doesn't say how. I couldn't figure it out.
- Figured out that I have to configure REST services and replace tokens in the advanced settings. Using getId.t2flow I got a new token and configured the REST services. Workflow worked OK.

General Comments

– What's the difference between getID.t2flow and getID.2.t2flow on http://opentox.informatik.tu-muenchen.de/trac/TUMOpenTox/wiki/Taverna#no1?

Conclusions and Recommendations

The online instructions should be detailed out quite a bit, and merged with the handout.

A.4 QPRF reporting

A.4.1 Review of Q-Edit by Roman Affentranger, 24 August 2011

OS: Windows 7. Java: Version 6 Update 26 (build 1.6.0_26-b03)

Purpose of Test

To carry out a first impressions evaluation of new version of QPRF, as well as of the updated documentation, to provide feedback to the developer.

Resources

Workshop Abstract: http://www.opentox.org/meet/opentox2011/opentox2011qprf

Q-Edit blog at http://opentox.ntua.gr/index.php/blog/78-q-edit

Download and start software

Used the download link at http://opentox.org/downloads: http://opentox.ntua.gr/files/Q-edit.jar

- Software starts up and presents empty window
- The Help Menu does only provide anything for "About…", but not for "Online Help", thus, I use the Q-Edit blog for guidance.
- The first page of the blog gives an introduction and provides some links to further information, as well as download links for Q-edit. That's nice.
- **PROBLEM**: at the bottom of the page, there are two "Prev" and "Next" links (see screenshot).

In fact, I hadn't seen the smaller text-links at first, and used the "Next >" button at the bottom-right. This one, however, links to http://opentox.ntua.gr/blog/77-ssl-certificates, a page about SSL certificates that only explain how to create them under Linux (thus scaring the vast majority of Windows users – many of which may not even realize that the instructions do not refer to Windows and who might try to "Open a terminal" under Windows and struggle with "mkdir", "cd", etc.). It doesn't say on the SSL page whether creating these certificates is necessary or just a feature making life with Q-edit nicer ;)
- Only by chance I found that the text-style link "Next >>" leads to somewhere else (http://opentox.ntua.gr/blog/78-q-edit?showall=&start=1).
- Unfortunately, the second paragraph already mentions again installing the SSL certificates (with a link to the corresponding blog page). No information what to do as a Windows user. I'm leaving this aside for now and continue (maybe a pointer that it's not strictly necessary would suffice while working out instructions for Windows users).
- The instructions for running Q-edit are again very much Linux-centered, and actually quite detailed, giving even command-line download instructions for Ubuntu users.
- Page 3 of the blog gives an overview over what one can do with Q-edit.
- Page 4 tells me to start by creating a new report. Did that…
- Page 5 explains how to load a compound. I used the given example, epicorticosterone. Worked fine. Tried to see if I still get something when having typos (epicorticoserne, epicorticosteron); Nothing found.

Compound Tab

- Noticed that one of the synonyms of epicorticosterone is "Molecule–1". Tried searching for "Molecule–1". Got Methanol. Methanol was depicted as a single "O":
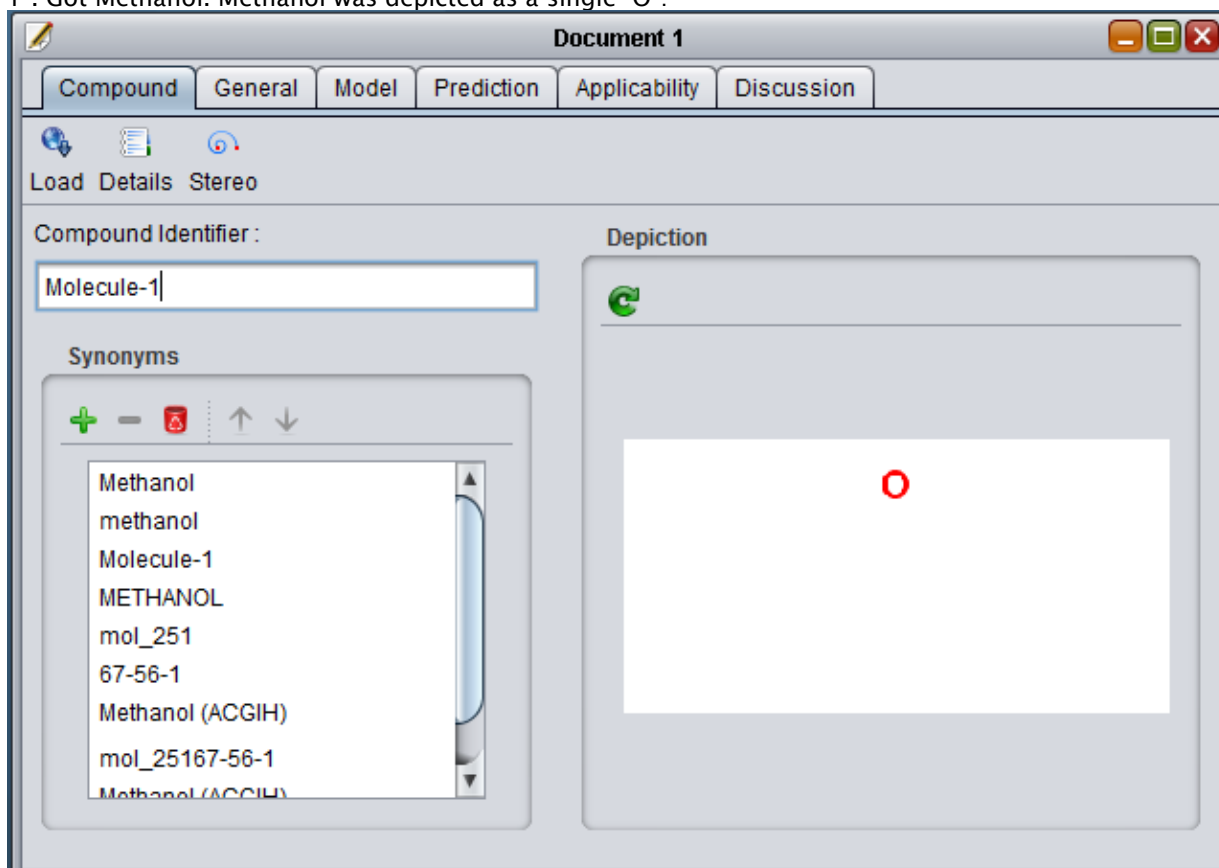
- Searching once more with Molecule–1 I got Meloxicam. Searching a third time with "Molecule–1" I got testosterone… Some synonyms seem a bit strange…
- Another strange synonym thing: "corticosterone" is given as a synonym for "epicorticosterone". However, the reverse is not true. Searching for "corticosterone" does not give any synonyms at all…
- Provided stereochemical information (bla bla).
- Removed synonyms, added one
- **GENERAL REMARK**: additional windows – e.g. the stereo one, or the one for adding synonyms – open up outside of the Q-edit window.
- Tested using the compound URI given in blog – worked OK

General Tab

- Adding authors: blog says the first two fields are mandatory. Maybe they could be marked with a "*" in the dialog box.
- Tried to fool the system by adding a space for my last name – didn't work ;)
- Tried adding myself a second time – didn't work, as mentioned in the blog.
- Setting the date: can't this be made such that by default it uses the current date taken from the computer (user can still change it, but for most cases that would probably be easier).
- Set the date to 1 March 2015
- Unclear what to do on the "Report" sub-tab of the "General" Tab.

Model Tab

- Chose at random a model from http://opentox.ntua.gr:8080/model (http://opentox.ntua.gr:8080/model/41f65c13–2bf9–4cb9–bb10–d63444377af6). A box pops up asking for my OpenTox credentials – which I give. Turns out I don't have access rights.

- Tried http://opentox.ntua.gr:8080/model/fffd62d8-e63b-4cfc-9b4a-c6ac441c78e9. This one loaded OK.
- Nice how one can get model information. "Find Online" button doesn't do anything.
- **NOTE** Aside: Incidentally, the model had not Unit for the prediction feature ☹
- **NOTE** Aside: the dependent feature of the model, caco2 (http://apps.ideaconsult.net:8080/ambit2/feature/22200), has "logP_{eff}" as unit. That's not a unit, is it? In Q-edit, I deleted the unit.
- **NICE** that one can edit information. However, this might also be a danger. Maybe, for these things that are all derived from a URI, a warning could be displayed when a user changes something…
- **UNCLEAR**: What should I add as "Date" and "Model Version"? Today's date? The date the model has been created? How can I know "Model Version"?

Prediction Tab

- Prediction Tab: prediction worked OK. However, the second text field is 1) too small for commenting, 2) unclear what should be entered there.
- Experimental value was downloaded OK. – Now I realized that the second text field is actually for the unit. This should be labeled.
- Obviously, the "Discussion" sub-tab of the Prediction tab is for discussing the predictions ;)

Applicability Tab

- The Applicability domain tab looks interesting. The blog, however, doesn't cover it. I wanted to test http://apps.ideaconsult.net:8080/ambit2/algorithm/pcaRanges, so I enter the URI and a name, but there's no "Run" or "Apply" button. The "Applicability Domain Result" has been a green check-mark before, and it stays one after that… **QUESTION**: what's the AppDomain tab for?
- Structural analogues retrieval takes a few seconds, but works fine. Also the compound info is nice.
- NOTE: one of the analogues for epicorticosterone is corticosterone. It's a different molecule! How come "corticosterone" is a synonym of epicorticosterone?
- Downloading experimental values: there are experimental values for two of the analogues.

GENERAL REMARK: the blog stops explaining here. What are the options for generating the report? Etc.

General things about Q-edit

- File -> Editor Options doesn't do anything
- Tools -> Options isn't doing anything
- Help -> Help Online isn't doing anything
- Help -> License Info isn't doing anything
- Used Tools -> Authentication to log in. It told me that Authentication completed. When using the "Log in" button, I'd expect a message that I'm already logged in, but I am just prompted for log in details again. It would be nice to have some indication whether one is currently logged in or not.
- I saved the report (both in Q-edit format, and as PDF). Then I changed somethings, then I clicked the "Exit the application" button. **NICE**: I'm asked if I really want to exit. Tested again what happens when I close the application without having changes in the report: I am not asked. SUGGESTION: since it seems dependent on whether or not there are unsaved changes, the popup could ask not only "Are you sure you want to exit the application", but might also mention that there are unsaved changes, and possibly even prompt for changes.

- When saving as PDF, the file type in the "saving browser" (don't know how that's called) isn't *.pdf, but "QPRF report".
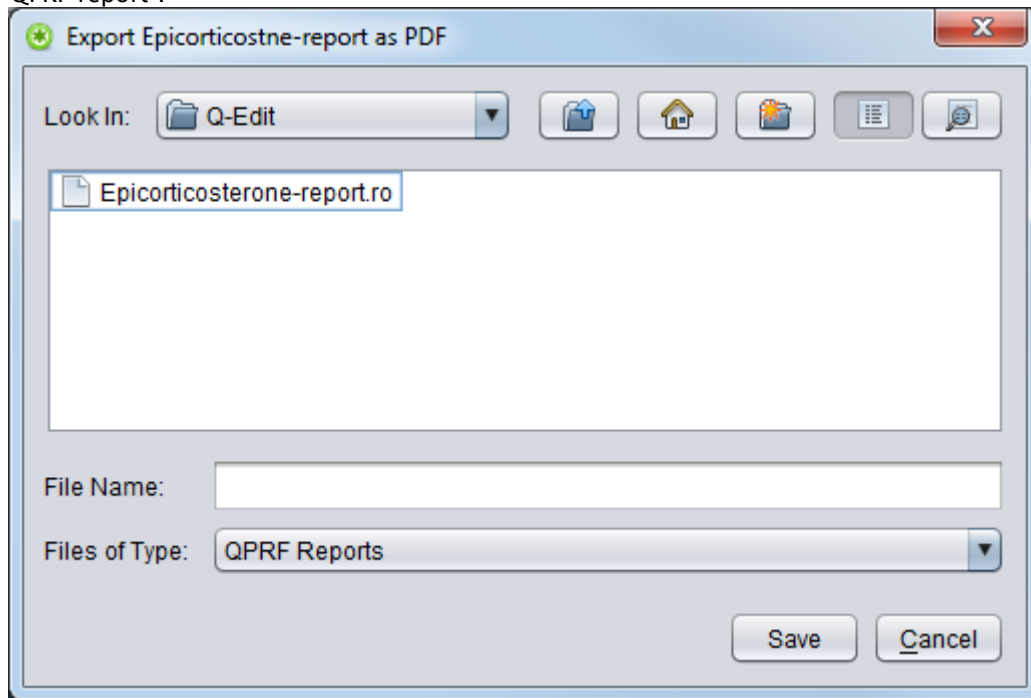
- When saving as PDF: when trying to overwrite an already existing PDF that was open (thus cannot be overwritten), I only get a small warning/error message at the bottom of the Q-edit window. However, the "Export" window closes. It's easy to think that the saving worked out OK. The warning should be made more prominent.

PDF report

- 3.1.a. Endpoint: where is that entered in Q-edit?
- 3.2.c. Reference to QMRF: where's that made?

Conclusions and Recommendations

Guidance is very helpful. It could, however, be extended a little. In particular, it should be made clear which instructions are for Linux, and which are for Windows. Windows instructions should be added. The blog-option should be revisited, as the paging is confusing.

Q-edit seems to work quite OK. However, it is not quite clear which fields result in which sections in the report.

The application has some buttons/menu items that are not functional. These are confusing.

A.5 CheS-Mapper

A.4.1 Review of CheS-Mapper by Barry Hardy, 2-6 August 2011

OS: Windows 7, 64 bit. IE8.

Description of Software Tested

Initial evaluation of the Ches Mapper software referenced in the OpenTox chapter in the AXLR8 2011 publication:

Scientific researchers in the field of chemoinformatics, are often overwhelmed by the size and the sheer complexity of chemical datasets. We have developed the 3D molecular viewer CheS-Mapper (Chemical Space Mapper) to enable the visualisation of complex chemical datasets used in predictive toxicology model building. Large datasets are divided into clusters of similar compounds and consequently arranged in 3D space, such that their spatial proximity reflects their chemical similarity (see Figure 5). This intuitively provides essential information to the user, while making the dataset more easily accessible and understandable. The different clustering approaches employed in the tool utilize common substructures as well as quantitative chemical descriptors of the compounds. These features can be highlighted within CheS-Mapper, which aids the user to better understand the underlying scientific knowledge. As a final function, the tool can also be used to select and export a specific part of a given dataset for further analysis.

Visualisation of Clustered Compounds with CheS-Mapper:
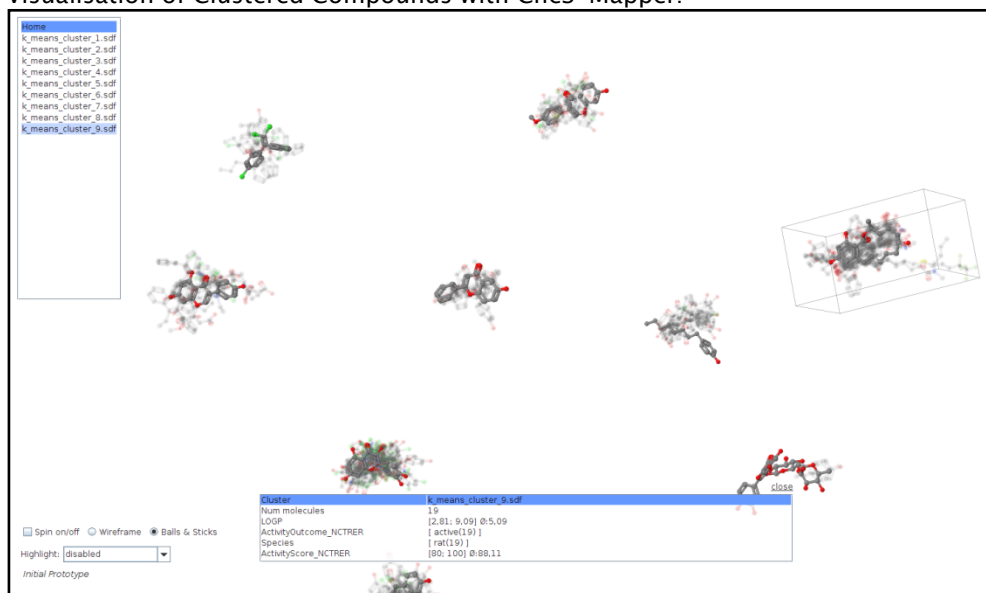
Purpose of Test

To carry out a first impressions evaluation of the CheS-Mapper software to provide feedback to the developer.

Poster http://vizbi.org/2011/Posters/Collection/?poster=C09

Web page with link to start the software:
http://opentox.informatik.uni-freiburg.de/ches-mapper/

As input you can use SDF-files and OpenTox Datasets from Ambit's dataset service. A nice dataset to start with could be this one (just copy and paste the URI into the applications dataset text field):
http://opentox.informatik.uni-freiburg.de/ches-mapper/data/caco2.sdf

Description of Test Results

First view of page:

   a)   Page does not yet provide explanations or guidance on what this SW is or what to do.

First Test:

1) I clicked on the Java web start and after a few seconds delay the application opened
2) Step 1: Load data step: It would be nice if a default set with abstract explanation/reference was available. No datasets were available under recently loaded datasets. I then entered the caco2 set of 100 molecules (sdf)
3) Step 2: Create 3D. I received the message that 3D cords were already available so I did not create any. First reaction: Why are some options in italics?
4) Step 3: Select features. It would be helpful if some documentation was available on what the features were that was in context, e.g. CDK feature definitions would appear on mouseover. I select caco2 and xlogP.
5) Step 4: Cluster data. I select the structural cluster structure web service of TUM. I am shown theta with a default value of 0.8. Theta is not defined.
6) Step 5: 3D Embedding. I go with the weka default. It is good that there is a description of this, but when I click on complete description I get a very wide bar that goes off the screen. It would be helpful if the parameter variation options were explained.
7) Step 6: Cluster alignment: I go with default of no alignment, and select Start Mapping.
8) I receive a message box which shows 25% Chemical Space Mapping, Building Cluster model with TUM algorithm web service. This is still showing the same message after 5 minutes, so I go and do other work in parallel. I check back after 20 minutes and get an error: Error 25% chemical space mapping. I abort. (Is it possible to pre-process for the workshop tutorial?)

Second Test:

I go through the steps in the first test but skip the clustering operation so that I can reach the visualization window. The visualisation window launches after a few seconds and I find myself in the middle of a rotating set of molecules, kind of looking like starships in a destroyed star fleet after the Borg have been though. They are in a plane so I assume I am seeing a two components separation based on the two properties I had selected (caco2, xlogP). It is not immediately clear what meaning axes directions in space have. I play around with the graphics options for a few minutes and am able to use a number of functions, change views etc. It is nice being able to see the properties associated with individual molecules. Some guidance at this point on options available to do some analysis would be useful. Trying it with the clustering next time will be nice to see.

Conclusions and Recommendations

My first experience in using the tool was quite good (e.g., going through the steps, visualizing molecules), but I was not able to do much with it (e.g., because clustering at TUM failed).

I need more help information available to continue testing. The clustering needs to be fixed so that it works.

A.4.1 Review of CheS-Mapper by Roman Affentranger, 18 August 2011

OS: Windows 7. Java: Version 6 Update 26 (build 1.6.0_26-b03)

Description of Software Tested

Initial evaluation of the version of CheS-Mapper downloaded 18 Aug 2011 from http://opentox.informatik.uni-freiburg.de/ches-mapper.

Purpose of Test

To carry out a first impressions evaluation of the software.

Resources

http://opentox.org/meet/opentox2011/opentox2011visualise

krystek.sdf downloaded from http://opentox.informatik.uni-freiburg.de/ches-mapper/data/krystek.sdf

1. Description of Test Results:

General

- There's no "Maximize", "Minimize" or "Restore" option for the CheS-Mapper?
- It doesn't have a taskbar-icon… Makes it hard to recover it when it's hidden behind a couple of windows.
- For methods offering a "complete description…" link, the pop-up window has no line-breaks, resulting in very wide windows.
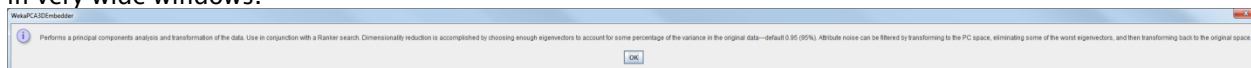
- What are the reasons for the severe limiations to <100 compounds (<50 with structural clustering?)?
- Parameters are not explained. Yet, if one changes them, the next time they're the same. Don't know anymore which were default parameters anymore. Maybe a "Reset defaults" for every method?

Loading a dataset

- When Opening a file and double-clicking a file in the file browser, the dialog box closes without selecting the file. It's common to select files by double-clicking, so this behavior should be corrected.

Create 3D structure

- No guidance/explanations about the differences of the three offered methods, neither in CheS-Mapper nor on the web.

Select Features

- Guidance missing
- Impact of selecting too many, too few?
- I selected all existing and computed features.

Cluster dataset

- No explanations for parameters
- Used structural clustering with default parameter of 0.4 for theta

3D Embedding

- No clue how the different methods differ, or what the parameters mean.
- Used WekaPCA3DEmbedder with default parameters.

3D Alignment

- MCS Aligner was grayed out, can be selected anyways, but get a message "OpenBabel comman obfit could not be found". Doesn't tell me what I can do about that…
- Selected "No cluster alignment" and started mapping.

Results

- "Zoom out" only works to some degree. At some point one cannot zoom out further.
- I don't know why I chose these 43 descriptors.

- Do the chosen descriptors have anything to do with the highlights?
- Is it possible to go back and change e.g. just the last setting? I could only exit the application, closing everything.

2. Second test

- Used http://opentox.informatik.uni-freiburg.de/ches-mapper/data/cristalli.sdf
- Had 3D structures generated
- Selected only few descriptors
- Used Hierarchical clustering
- Results were a bit more illustrative, as I got more than just one cluster ;)
- Found the way to go back to change settings... (File -> New dataset/mapping). Possibly rename this?
- Tried a different clustering (EM). Worked fine, but could not rotate anymore...

3. Third test

- Used http://opentox.informatik.uni-freiburg.de/ches-mapper/data/NCTRER_v4b_232_15Feb2008.ob3d.sdf
- Played around with different descriptor selections (always without clustering). The third time rotation didn't work anymore, then also highlighting stopped working, then also zooming, ...

4. Fourth test

- 696 compounds, 8 features, Kmeans, WekaPCA3D -> OK
- 696 compounds, 8 features, Hierarchical clustering, WekaPCA3D -> OK
- 2123 compounds, 3 features, Cobweb, WekaPCA3D -> **ERROR (Java Heap Space)** (SD file: https://docs.google.com/leaf?id=0B4p1CPY3IHzvNDY5YjA2MzUtZGU1YS00NzcyLWE1YjQtOGU0MTJlNDk0MGQ3&hl=en_US)
- 2123 compounds, 3 features, No clustering, WekaPCA3D -> **ERROR (Java Heap Space)** (SD file: https://docs.google.com/leaf?id=0B4p1CPY3IHzvNDY5YjA2MzUtZGU1YS00NzcyLWE1YjQtOGU0MTJlNDk0MGQ3&hl=en_US)
- 1524 compounds, 7 features, No clustering, WekaPCA3D -> **ERROR (Java Heap Space)** (SD file: https://docs.google.com/leaf?id=0B4p1CPY3IHzvMGE3NzM4N2YtY2I2Mi00NjIxLTk0OTgtMjY5ODU0Y2I2NmFi&hl=en_US)

Conclusions and Recommendations

It's difficult to test the application without knowing how it is working, what it is supposed to do, etc.

This Appendix summarises the evaluation results for the following 9 exercises which were run at the hands-on OpenTox workshop in Munich 9 August 2011:

A. Use ToxPredict to obtain REACH-relevant information on compounds
David Gallagher, Barry Hardy (More Information ...)

B. Use ToxCreate to build a (Q)SAR model
Christoph Helma, David Vorgrimmler (More Information ...)

C. Validation and Reporting of Models
Andreas Maunz, Harry Sarimveis (More Information ...)

D. Create and run Workflows across OpenTox services
Joerg Wicker, Fabian Buchwald (More Information ...)

E. Generate Predictions and their QPRF Reports
Pantelis Sopasakis, Sunil Chawla (More Information ...)

F. Predict Metabolites of Compounds
Patrik Rydberg, Nina Jeliazkova (More Information ...)

G. Apply Bioclipse and OpenTox to modify Chemical Structure to remove Toxicity
Ola Spjuth (More Information ...)

H. Build a Weight of Evidence for Drug Candidate Molecules
Roman Affentranger, Indira Ghosh (More Information ...)

I. Visualise and Explore Chemical Feature Space of a Dataset
Martin Gütlein, Andreas Karwath (More Information ...)

OpenTox Munich workshop feedback summary (9th August 2011)

The summaries below were transcribed from the recorded web responses, and included most of the comments.

A. ToxPredict

What is your perception of the purpose(s) of this software? (summarise)

In silico toxicity profiling for single compounds, present all available data
If it is aimed at someone not so skilled in statistics – leave the multiple plotting etc and produce values that are interesting to the enduser together with an explanation what it means in this particular case. For example, a plot explaining how the value of the confidence threshold was selected, and what it means for MSE.   Explanations to what the terms means.
Looks promising but needs to widen the scope
Generation of toxicity report for chemical compounds based on various toxicity models
Easy access to predictions from available in silico models related to toxicology.
I think it is very useable for prediction of different properties for chemical compounds.
Easy to use, covers validation and the prediction is quite well explained

How do you rate the software overall in terms of its scientific interest:

excellent	very good	average	needs improvement	poor
3	5	2	0	0

How do you rate the software overall in terms of its ease of use:

3	8	0	0	0

How do you rate the software overall in terms of its technical merit:

3	3	2	0	0

How well does the software convey its purpose & functionality?

2	2	1	3	0

How understandable are the results and information generated?

1	3	3	3	0

How do you rate the look and feel of the interface?

2	6	2	0	0

How do you rate how easy it was to understand what you were doing?

3	4	3	0	0

How do you rate how the system responded to your explorations?

1	4	3	1	0

How do your rate the help and guidance provided by the system?

2	3	3	2	0

How well were you able to evaluate the information provided?

1	3	3	2	0

How reliable was the information provided?

1	4	4	0	0

How relevant was the information provided?

2	2	4	1	0

How adequate or useful was the information provided?

2	6	2	0	0

How complete was the information provided?

1	3	4	1	0

Are there any issues you found irritating or annoying?

The software did crash a few times whilst using it in this workshop, but this did not bother me too much as Toxpredict is still relatively new.
QSAR SRC KOWWIN fingerprints AD Don"t understand the meaning of this No alerts for the micronucleus assay (Class II)YES Ambigous in english, what does this mean? Are there any..? web service for the Caco-2 Unit missing Confidence 0.16844466 Scale? 8 Digits?? Are you confident this is the exactness? IST Fish Toxicity model (LC50_mmol)   Calculate Confidence 0.16844466 LC50_mmol0.59865606 Acute or chronic toxicity test?   ECOSAR LC50 fish Calculate F96h-LogLC50mmol/L(expt_provided in ECOSAR manual) Mode of action of the predicition - narcosis, active reagent.. ?
More links to documentation of methods, data sets etc.

What did you particularly like?

Good attentive teachers. Good worked examples.
Simplicity, but very incomplete and non-standardized returns
Easy to use.
The simplicity and user friendly interface

Please indicate any specific comments, needs or questions you have.

None reported

B. ToxCreate

What is your perception of the purpose(s) of this software? (summarise)

Based on the requirement of a particular project, toxcreate can certainly help in getting some baseline information to better inform the predictive aspects of chemical toxicology.
Creating models to predict toxicity, inspecting models, getting validation results, making predictions
Easy, rapid tool to get toxicity endpoint predictions for non-experts in machine learning
Very well designed piece of software. I would have to get more experience with it, to be able to evaluate it. Some statistical issues may be present, such as the 0-1 based confidence measure. It may be worthwhile to add probabilistic measures to the technology
I intend to develop and validate QSAR models
to create models to predict biological activity
Create models for predicting properties of compounds

How do you rate the software overall in terms of its scientific interest:

excellent	very good	average	needs improvement	poor
3	3	3	0	0

How do you rate the software overall in terms of its ease of use:

5	6	0	0	0

How do you rate the software overall in terms of its technical merit:

2	4	4	0	0

How well does the software convey its purpose & functionality?

3	4	2	0	0

How understandable are the results and information generated?

5	3	3	0	0

How do you rate the look and feel of the interface?

2	5	4	0	0

How do you rate how easy it was to understand what you were doing?

4	5	1	1	0

How do you rate how the system responded to your explorations?

1	3	4	0	0

How do your rate the help and guidance provided by the system?

2	6	1	0	0

How well were you able to evaluate the information provided?

2	4	3	0	0

How reliable was the information provided?

1	5	3	1	0

How relevant was the information provided?

5	5	3	0	0

How adequate or useful was the information provided?

2	4	2	0	0

How complete was the information provided?

3	4	1	1	0

Are there any issues you found irritating or annoying?

None cited

What did you particularly like?

I like the straightforward user interface and relatively uniform design of the different modules
Upload format flexibility
hands-on demonstration, running real examples
Ease of use

Please indicate any specific comments, needs or questions you have.

I personally would like to study the lazar methodology and see if I could contribute to it with probabilistic methods
Sound expert report templates for good, average and poor predictions would help for regulatory work.
Would be nice to see such services deployed at more locations

C. Validation & Reporting

What is your perception of the purpose(s) of this software? (summarise)

Software for Validation and reporting of model performance.
To carry out validation procedures of qsar models and generate reports to summarize the results
Positive. User has a lot of flexibility to use this software. I was primarily interested in the application of the type of information from the perspective of human health risk assessment.

How do you rate the software overall in terms of its scientific interest:

excellent	very good	average	needs improvement	poor
2	5	1	0	0

How do you rate the software overall in terms of its ease of use:

0	2	5	1	0

How do you rate the software overall in terms of its technical merit:

3	5	0	0	0

How well does the software convey its purpose & functionality?

0	9	1	0	0

How understandable are the results and information generated?

1	4	3	0	0

How do you rate the look and feel of the interface?

0	4	3	1	0

How do you rate how easy it was to understand what you were doing?

0	5	3	0	0

How do you rate how the system responded to your explorations?

1	3	3	1	0

How do your rate the help and guidance provided by the system?

0	3	4	0	0

How well were you able to evaluate the information provided?

1	4	2	0	0

How reliable was the information provided?

5	3	4	0	0

How relevant was the information provided?

0	6	2	0	0

How adequate or useful was the information provided?

1	0	0	0	0

How complete was the information provided?

1	0	0	0	0

Are there any issues you found irritating or annoying?

Very limited model output, no confidence bounds, or possibly plots related to the validation method(s)
some error messages while running algorithms and accessing servers
The OpenTox framework is new to me and hence I need further information to be able to use it

What did you particularly like?

machine learning approach
reports are very informative
Everything is an uri

potential applications in the future and transparent nature of the system

Please indicate any specific comments, needs or questions you have.

add more chemical/toxicological knowledge
Keep going.
would like to discuss potential development and applications in the context of human health risk assessments.

D. Workflows

What is your perception of the purpose(s) of this software? (summarise)

To facilitate usage of the REST services in a graphical environment.
Great flexibility in developing frames for managing datasets along with results.

How do you rate the software overall in terms of its scientific interest:

excellent	very good	average	needs improvement	poor
1	2	0	0	0

How do you rate the software overall in terms of its ease of use:

1	1	1	1	0

How do you rate the software overall in terms of its technical merit:

0	1	1	0	0

How well does the software convey its purpose & functionality?

0	0	2	0	0

How understandable are the results and information generated?

0	1	1	0	0

How do you rate the look and feel of the interface?

1	1	1	0	0

How do you rate how easy it was to understand what you were doing?

0	0	4	0	0

How do you rate how the system responded to your explorations?

0	2	0	0	0

How do your rate the help and guidance provided by the system?

0	1	3	0	0

How well were you able to evaluate the information provided?

0	0	2	0	0

How reliable was the information provided?

1	0	0	0	0

How relevant was the information provided?

1	0	0	0	0

How adequate or useful was the information provided?

1	0	0	0	0

How complete was the information provided?

1	0	0	0	0

Are there any issues you found irritating or annoying?
None reported

What did you particularly like?

The chance of mixing dataset

Please indicate any specific comments, needs or questions you have.
None reported

E. Predictions and QPRF

Only one partial response

F. Predict Metabolites

What is your perception of the purpose(s) of this software? (summarise)

To predict the possible sites of metabolism and the type of prediction expected. The prediction is based on QM model that is not trained with any experimental data.
predict p450 sites of metabolism and generate metabolites
f. integration into existing workflows
generally it work well. for certain class of compounds, the information may need to be validated.

How do you rate the software overall in terms of its scientific interest:

excellent	very good	average	needs improvement	poor
2	3	0	0	0

How do you rate the software overall in terms of its ease of use:

3	2	0	0	0

Are there any issues you found irritating or annoying?
None reported

What did you particularly like?

ease of use
original approach
very user friendly approach
ease of use of toxtree

Please indicate any specific comments, needs or questions you have.

would be nice to apply the same methodology to other targets
would be interested to discuss potential development where information can be updated regularly for understanding chemical toxicology information and validation of predicted information

H. Weight of Evidence

What is your perception of the purpose(s) of this software? (summarise)

REACH related prediction, proof of concept of modelling framework

How do you rate the software overall in terms of its scientific interest:

excellent	very good	average	needs improvement	poor
1	4	0	0	0

How do you rate the software overall in terms of its ease of use:

1	1	2	1	0

How do you rate the software overall in terms of its technical merit:

1	3	0	1	0

How well does the software convey its purpose & functionality?

0	3	2	0	0

How understandable are the results and information generated?

2	1	1	1	0

How do you rate the look and feel of the interface?

0	2	1	1	0

How do you rate how easy it was to understand what you were doing?

0	4	1	0	0

How do you rate how the system responded to your explorations?

0	3	0	1	0

How do your rate the help and guidance provided by the system?

0	5	0	0	0

How well were you able to evaluate the information provided?

1	3	1	0	0

How reliable was the information provided?

0	4	3	0	0

How relevant was the information provided?

1	1	3	0	0

How adequate or useful was the information provided?

1	2	2	0	0

How complete was the information provided?

0	3	1	1	0

Are there any issues you found irritating or annoying?

Simple and detailed information on the concepts like e.g. what is a feature in this context technically etc.

What did you particularly like?

The way results can be displayed, i.e. in PDF format
Good teachers. Excellent written instructions.
Introduction, good documentation

Please indicate any specific comments, needs or questions you have.

None reported

I. Visualization

What is your perception of the purpose(s) of this software? (summarise)

Nice tool to vizualize the chemical space. See similarities and differences.
Visualisation of clustering results
The purpose is to provide a visualization of chemical space which in turn will allow the identification of features that confer a biological property to a chemical
Obtain a nice view of chemical space that is easily comprehensible
to cluster compounds in descriptor space

How do you rate the software overall in terms of its scientific interest:

excellent	very good	average	needs improvement	poor
2	2	1	0	0

How do you rate the software overall in terms of its ease of use:

1	3	0	0	0

How do you rate the software overall in terms of its technical merit:

1	2	2	0	0

How well does the software convey its purpose & functionality?

2	3	0	0	0

How understandable are the results and information generated?

2	2	1	0	0

How do you rate the look and feel of the interface?

5	1	0	0	0

How do you rate how easy it was to understand what you were doing?

0	5	0	0	0

How do you rate how the system responded to your explorations?

1	1	2	1	0

How do your rate the help and guidance provided by the system?

0	5	1	0	0

How well were you able to evaluate the information provided?

2	2	0	0	0

How reliable was the information provided?

1	2	0	0	0

How relevant was the information provided?

1	2	1	0	0

How adequate or useful was the information provided?

0	2	2	0	0

How complete was the information provided?

0	3	1	0	0

Are there any issues you found irritating or annoying?

Failed on most of the datasets tested, testing needed
My SDFile could not be read correctly so I cannot evaluate the last questions. It's probably a matter of reading the sdf, however, the bug has been reported and will be looked into.
need to be able to zoom out further

What did you particularly like?

Quick help for installation issues. nicely prepared data.
Clear user interface, easily usable
First tool that I have come across that allows easy visualization of chemical space

The design and easiness of use
relating to chemical structure on screen ability to calculate descriptors as well as input of user descriptor

Please indicate any specific comments, needs or questions you have.

Consider adding structural alert features
would like ability to show overlap between different data sets

Appendix C Models built in OpenTox on REACH endpoints

This appendix lists a representative subset of the models created by OpenTox partners on REACH endpoints. The list is not comprehensive, as this is an ongoing effort and several models are in the stage of still being optimized.

C.1 Carcinogenicity

Dataset: ISSCAN v3a carcinogenicity dataset (http://www.epa.gov/comptox/dsstox/sdf_isscan_external.html, http://www.ncbi.nlm.nih.gov/pubmed/18469376)

OpenTox URI of dataset: http://apps.ideaconsult.net:8080/ambit2/dataset/9

Description of Dataset: General classification scheme: 3 = "carcinogen/mutagen", 2= "equivocal", 1 = "non-carcinogen/non-mutagen".

Model URI: http://lazar-services.in-silico.ch/model/270
Validation URI: http://lazar-services.in-silico.ch/validation/report/crossvalidation/259
Type: Classification model
Created by partner: IST
Algorithm used: Lazar, BBRC tree-shaped fragments
Name in ToxPredict: –

C.2 Carcinogenicity

Dataset: CPDBAS: Carcinogenic Potency Database Summary Tables – All Species (http://www.epa.gov/ncct/dsstox/sdf_cpdbas.html)

OpenTox URI of dataset: Several datasets. See model description below.

Description of Dataset: Contains carcinogenic potencies for several species (mouse, rat, hamster, single cell, multi cell)

Model URI: http://lazar-services.in-silico.ch/model/193
Dataset URI: http://lazar-service.in-silico.ch/dataset/2241
Validation URI: http://lazar-services.in-silico.ch/validation/report/crossvalidation/22
Type: Classification model
Created by partner: IST
Algorithm used: Lazar
Name in ToxPredict: IST DSSTox Carcinogenic Potency DBS Hamster

Model URI: http://lazar-services.in-silico.ch/model/204
Dataset URI: http://lazar-service.in-silico.ch/dataset/2283
Validation URI: http://lazar-services.in-silico.ch/validation/report/crossvalidation/23
Type: Classification model
Created by partner: IST
Algorithm used: Lazar
Name in ToxPredict: IST DSSTox Carcinogenic Potency DBS Mouse

Model URI: http://lazar-services.in-silico.ch/model/215
Dataset URI: http://lazar-service.in-silico.ch/dataset/2325
Validation URI: http://lazar-services.in-silico.ch/validation/report/crossvalidation/24
Type: Classification model

Created by partner: IST
Algorithm used: Lazar
Name in ToxPredict: IST DSSTox Carcinogenic Potency DBS MultiCellCall

Model URI: http://lazar-services.in-silico.ch/model/237
Dataset URI: http://lazar-service.in-silico.ch/dataset/2409
Validation URI: http://lazar-services.in-silico.ch/validation/report/crossvalidation/26
Type: Classification model
Created by partner: IST
Algorithm used: Lazar
Name in ToxPredict: IST DSSTox Carcinogenic Potency DBS Rat

Model URI: http://lazar-services.in-silico.ch/model/248
Dataset URI: http://lazar-service.in-silico.ch/dataset/2251
Validation URI: http://lazar-services.in-silico.ch/validation/report/crossvalidation/27
Type: Classification model
Created by partner: IST
Algorithm used: Lazar
Name in ToxPredict: IST DSSTox Carcinogenic Potency DBS SingleCellCall

C.3 Salmonella Mutagenicity

Dataset: ISSCAN mutagenicity dataset (http://www.epa.gov/comptox/dsstox/sdf_isscan_external.html, http://potency.berkeley.edu/cpdb.html):

OpenTox URI of dataset: http://apps.ideaconsult.net:8080/ambit2/dataset/9

Description of Dataset: 403 active, 457 inactive compounds.

Model URI: http://lazar-services.in-silico.ch/model/226
Validation URI: http://lazar-services.in-silico.ch/validation/report/crossvalidation/25
Type: Classification model
Created by partner: IST
Algorithm used: Lazar, BBRC tree-shaped fragments
Name in ToxPredict: IST DSSTox Carcinogenic Potency DBS Mutagenicity

C.4 Fish/Aquatic Toxicity

Dataset: EPA DSSTox Fathead Minnow Dataset (http://www.epa.gov/ncct/dsstox/sdf_epafhm.html).

OpenTox URI of dataset: http://apps.ideaconsult.net:8080/ambit2/dataset/12

Description of Dataset: 617 compounds, LC50 data in mM. ActivityOutcome_EPAFHM: 564 active, 50 inactive, 3 inconclusive.

Model URI: http://lazar-services.in-silico.ch/model/259
Validation URI: http://lazar-services.in-silico.ch/validation/report/crossvalidation/28
Type: Regression model
Created by partner: IST
Algorithm used: Lazar, BBRC path-shaped fragments
Name in ToxPredict: IST EPA v4b Fathead Minnow Acute Toxicity (LC50_mmol)

Model URI: http://195.178.207.160/opentox/model/makeSCR_457
Validation URI: http://opentox.informatik.uni-freiburg.de/validation/1380
Type: Regression model
Created by partner: IBMC
Algorithm used: SCR
Name in ToxPredict: EPAFM LC50 fish OpenTox model created with SCR regression model

Appendix D Use ToxPredict to obtain REACH-relevant information on compounds

Facilitated by David Gallagher and Barry Hardy

Objectives

To demonstrate and evaluate a prototype OpenTox application (ToxPredict) that accepts chemical structures and names as input and automatically generates a toxicity report based on various pre-calibrated toxicity models. We will examine the relevance and usefulness of the information provided for potential REACH reporting purposes.

Description of ToxPredict

ToxPredict is a web-based interface for predicting toxicity of individual or sets of chemicals. Users can either search for a compound in the OpenTox prototype database, which currently includes quality labeled data for 243,648 chemicals (including all chemicals currently registered with REACH) grouped in 25,170 datasets, or upload their own chemical structure in the SDF format. It runs the selected calculations automatically using a collection of distributed computational services. ToxPredict currently includes 24 validated models addressing 11 different endpoints, and is extensible.

Prerequisites: A computer with browser (Internet Explorer or Firefox) and fast internet connection.

Note: any compounds not found in the OpenTox database, must be uploaded as SDF or MOL files.

Summary of Activities

In this workshop activity we will

a) Review the Information Gathering Requirements of REACH

b) Work through examples using the OpenTox-based application ToxPredict to discover and discuss what REACH-relevant information OpenTox currently provides

c) Provide time for users to work on their own examples (You may for example bring an sdf file with you for a compound or a library of ca. 5-100 molecules or you can draw your own structure and submit it for predictions)

d) Discuss requirements and features you would like to see included in the application for your future work.

e) Complete an evaluation providing your feedback.

Exercise Sequence

1) Login to Toxpredict (www.toxpredict.org) using userid "guest" and password "guest".

2) Example 1: Type the name Benzene into the query text box, select the "Auto Detect" option, and click "Search".

3) Check that benzene and its associated information page has been returned to you. Click the "Run All" button on this new page.

4) Take a few minutes to review the toxicity-related information presented to you.

5) Focus in on the IST Kazius-Bursi Salmonella mutagenicity classification model. Detailed model validation information can be found linked into the Browse Models area of ToxPredict. (Or is also available at http://lazar-services.in-silico.ch/validation/report/crossvalidation/6) What does the prediction mean? How reliable are the results?

6) Example 2: Try searching for nitrosamines (e.g., enter the SMILES CCN(CC)N=O in query box, select substructure match) and click "Search".

7) View structures presented to you, select one of interest, and Click the "Run All" button.

8) Take a few minutes to review the toxicity-related information presented to you.

9) Focus in on the IST EPA v4b Fathead Minnow Acute (Fish) Toxicity regression model. Detailed model validation information can be found linked into the Browse Models area of ToxPredict. (Or is also available at http://lazar-services.in-silico.ch/validation/report/crossvalidation/28) What does the prediction mean? How reliable are the results?

10) Example 3: Repeat the procedure for the dioxin 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD), [SMILES = Clc2cc1Oc3c(Oc1cc2Cl)cc(Cl)c(Cl)c3]

11) Example 4: Try out your own examples and discuss results.

12) Complete the evaluation form following this link: Evaluation Form