Powered by:

ToxOtis-core

ToxOtis-DB

JAQPOT Web Services

An Implementation of the OpenTox Application Programming Interface for On-line Prediction of Toxicological Properties

Pantelis Sopasakis[a] and Haralambos Sarimveis[b]

[a] IMT Alti Studi Lucca, Piazza San Ponziano, 6, 55100, Lucca, Italy. E-mail: p.sopasakis@imtlucca.it
[b] National Technical University of Athens, School of Chemical Engineering, Heroon Politechniou, 9, 15780, Zografou Campus, Athens, Email: hsarimv@central.ntua.gr

About

The landscape of predictive toxicology has utterly changed after the completion of the FP7-funded EU research project OpenTox [1] which established new standards for systems of distributed toxicological computation. JAQPOT is a bundle of web services, compliant to the OpenTox Application Programming Interface [2] and provides access to Machine Learning algorithms and data processing routines particularly tailored to the needs of predictive toxicology. End-users may exploit its functionalities either through web interfaces or programmatically. Toxicologists, biologists and other scientists can make the best out of their data using a transparent, modular and interoperable computational platform on the web. Confidential in-house data can also be integrated securely by means of an access control infrastructure based on Single Sign-On.

JAQPOT is available online at http://opentox.ntua.gr:8080 and is accessible through its REST interface, or can be downloaded from https://github.com/alphaville/jaqpot3 and installed locally on one's LAN following the instruction found at http://opentox.ntua.gr/blog/14-standalone-installation-jaqpot3.

Machine Learning Algorithms

JAQPOT provides access to a number of algorithms for model training and data preprocessing. According to the OpenTox API v1.2 algorithms interface.

Description	Method	URI	Result
Get a list of the URIs of all available algorithms	GET	`/algorithm`	List of all algorithm URIs or RDF representation of specific types in presence of a query parameter.
Get the ontological representation of an algorithm	GET	`/algorithm/{id}`	Algorithm representation in one of the supported MIME types.
Apply the algorithm	POST	`/algorithm/{id}`	The result is either a model URI, a dataset URI or a feature URI depending on the specifications of the algorithm. Long-running processes return the URI of a Task.

This is an excerpt from the ontological representation of the SVM algorithm of JAQPOT:

```
<http://opentox.ntua.gr:8080/algorithm/svm>
  a       ota:EagerLearning , ot:Algorithm , ota:Regression , ota:SingleTarget ;
  rdfs:comment "Representation automatically generated by ToxOtis."^^xsd:string;
  dc:date "Mon Apr 29 16:50:25 EEST 2013"^^xsd:dateTime ;
  dc:description "Algorithm for training …"^^xsd:string ;
  dc:identifier "http://opentox.ntua.gr:8080/algorithm/svm"^^xsd:anyURI ;
  dc:publisher "http://opentox.ntua.gr:8080/"^^xsd:string ;
```

The tuning parameters of an algorithm (e.g. the gamma parameter of an SVM algorithm) have an ontological representation as well. Here is an example:

```
<http://opentox.ntua.gr:8080/prm/svm_gamma>
  a       ot:Parameter ;
  rdfs:comment "Only strictly positive values are acceptable"^^xsd:string ;
  dc:description "Gamma Parameter for the SVM kernel"^^xsd:string ;
  dc:identifier "http://opentox.ntua.gr:8080/prm/svm_gamma"^^xsd:anyURI ;
  dc:title "gamma"^^xsd:string ;
  ot:paramScope "OPTIONAL"^^xsd:string ;
  ot:paramValue "1.5"^^xsd:double .
```

The following algorithms are available in JAQPOT:

- **Fast RBF Neural Networks**: The algorithm is based on the subtractive clustering technique and has a number of advantages compared to the traditional learning algorithms including faster training times and more accurate predictions. Due to these advantages the method proves suitable for developing models for complex nonlinear systems [3].
- **Partial Least Squares**: PLS is a standard, widely used supervised algorithm for dimension reduction on datasets.
- **Scaling**: This web service is intended to scale the numeric values of an OpenTox dataset within a specified range.If not otherwise specified by the client, this range is assumed to be [-1,1]. Scaling is necessary for algorithms like SVM and Neural Networks as it substantially improves the accuracy of the trained models.
- **Support Vector Machines**
- **Missing Value Replacer:** Replaces the missing values in a dataset using the means-and-modes technique.
- **Leverage:** Algorithm for the estimation of the Domain of Applicability of a model.
- **Multiple Linear Regression**

Train, Evaluate and Use (Q)SAR Models in JAQPOT

Before the users can use any of the services of JAQPOT they should authenticate their credentials against the SSO server of OpenTox and receive an authentication token: Here is an example cURL request:

```
curl -X POST https://opensso.in-silico.ch/opensso/identity/
authenticate?uri=service=openldap -k -d username=guest -d
password=guest
```

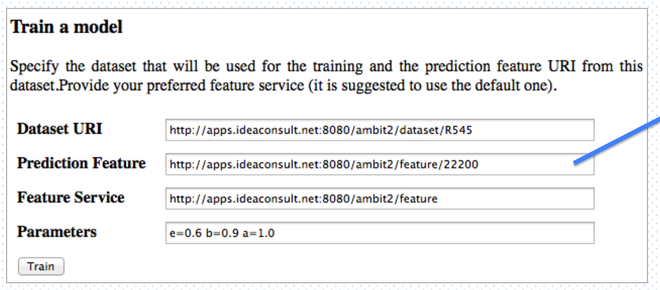
This request will return a token to the client (if the provided credentials are correct). For instance:

```
token.id=AQIC5wM2LY4Sf…b79U8OjM.*AAJTSQACMDE.*
```

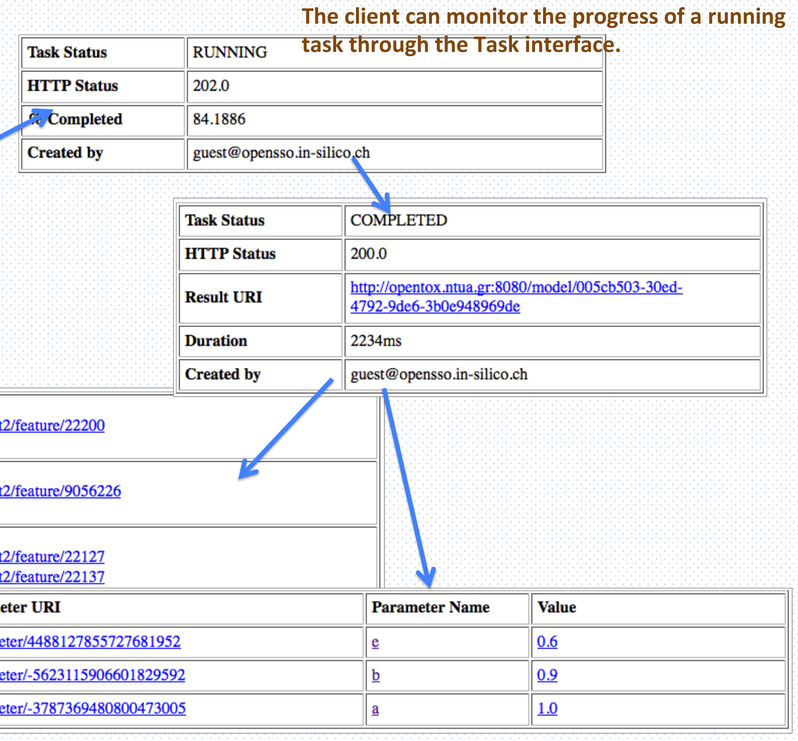
This token is passed to the algorithm service using the Header `subjectid` along with the tuning parameters of the algorithm, the URI of the dataset and the URI of the target feature.

```
curl -X POST http://opentox.ntua.gr:8080/algorithm/fastRbfNn
-d dataset_uri=http://apps.ideaconsult.net:8080/ambit2/dataset/R545
-d  prediction_feature=http://apps.ideaconsult.net:8080/ambit2/
feature/22200 -d a=1.0 -d b=0.86 -d e=0.6
-H "subjectid:AQIC5wM2L…CMDE.*" -H Accept:text/uri-list
```

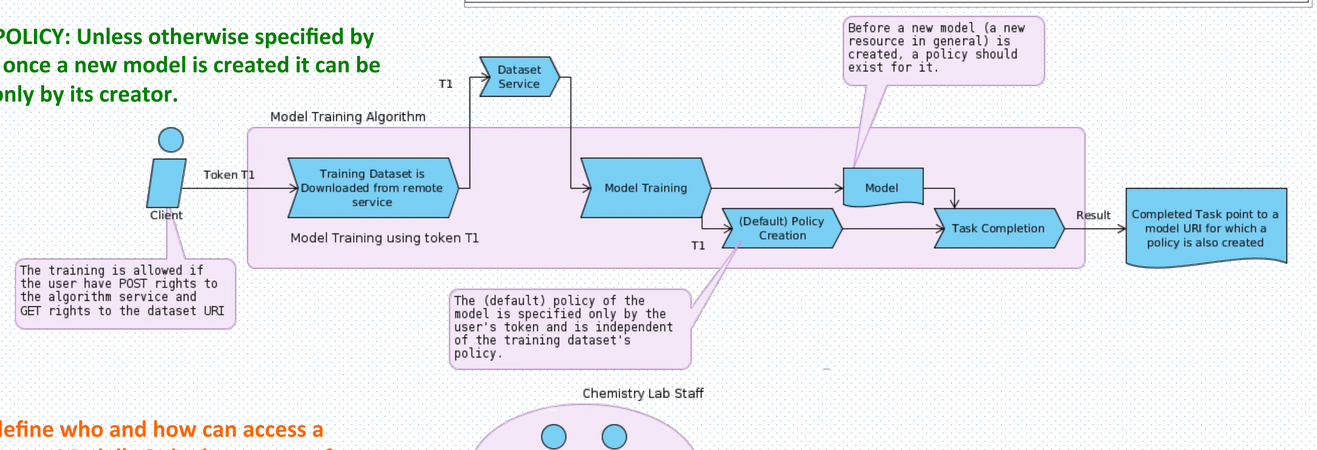
This will return the URI of a task which upon completion will return the URI of a model.



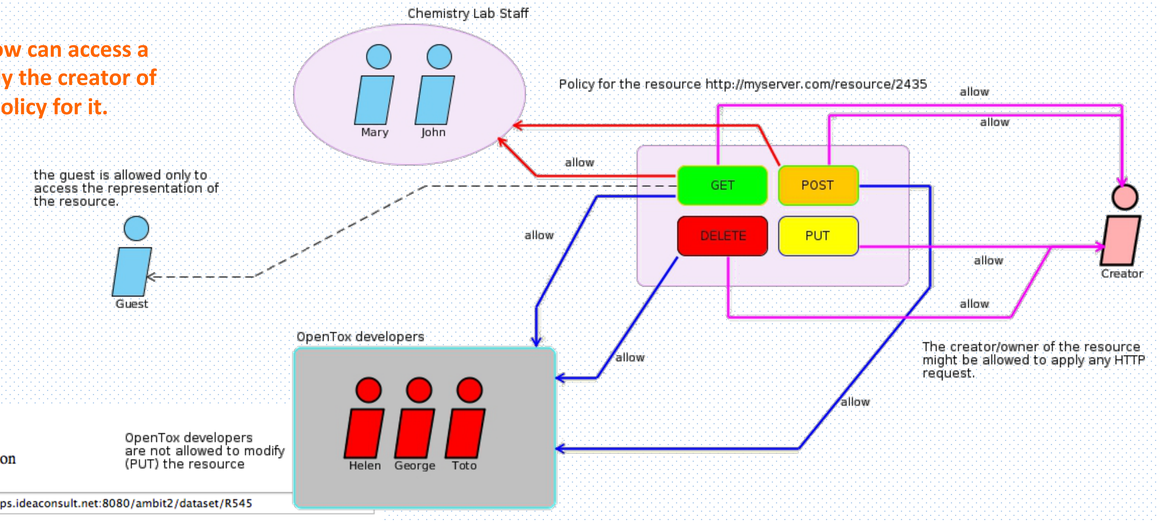
A simple HTML interface has also been setup for easier consumption of the web services.

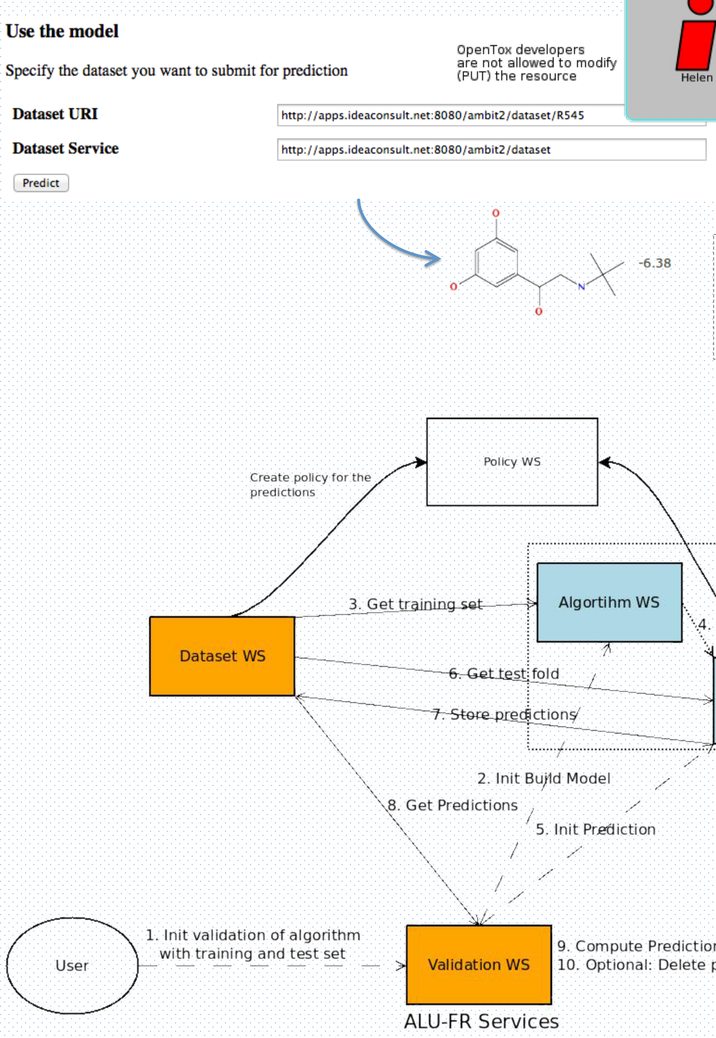


DEFAULT POLICY: Unless otherwise specified by the client, once a new model is created it can be accessed only by its creator.



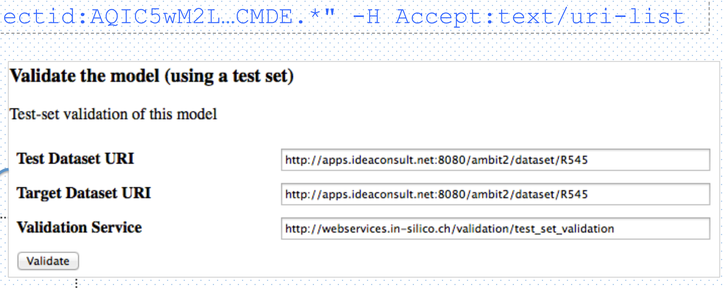
POLICIES define who and how can access a resource (e.g. a Model). Only the creator of a resource can modify the policy for it.

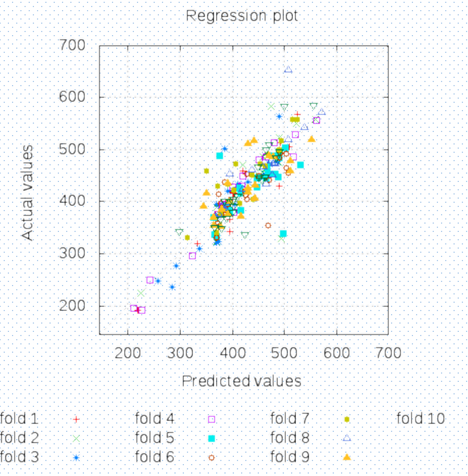




Use a (Q)SAR model to perform a prediction (This will create a new dataset were the estimated properties and the corresponding experimental values – if any – are stored):

```
curl -X POST opentox.ntua.gr:8080/model/005cb503
-30ed-4792-9de6-3b0e948969de-d dataset_uri=
apps.ideaconsult.net:8080/ambit2/dataset/R545 -H
"subjectid:AQIC5wM2L…CMDE.*" -H Accept:text/uri-list
```



Acknowledgement

This work was financially supported by **OpenTox** - An Open Source Predictive Toxicology Framework, http://www.opentox.org, is funded under the EU Seventh Framework Program: HEALTH-2007-1.3-3 Promotion, development, validation, acceptance and implementation of QSARs (Quantitative Structure-Activity Relationships) for toxicology, Project Reference Number Health-F5-2008-200787 (2008-2011).

References

1. Hardy, B. *et al.* 2010. Collaborative development of predictive toxicology applications. *Journal of Cheminformatics*, **2**, 1 – 29.
2. The OpenTox API ver. 1.2 – Technical Specifications: Available online at http://opentox.org/dev/apis/api-1.2 (accessed on 12 June, 2013).
3. H. Sarimveis, A. Alexandridis and G. Bafas 2003., A fast training algorithm for RBF networks based on subtractive clustering, Neurocomputing, 52, 501 – 505.
4. N. Jeliazkova and V. Jeliazkov 2011, AMBIT RESTful Web Services: An implementation of the OpenTox Application Programming Itnerface, *Journal of Cheminformatics*, **3**, 1 – 18.