GRIB CADD

Open PHACTS — Open Pharmacological Space

eTOX

http://cadd.imim.es

USE OF ADVANCED KNOWLEDGE MANAGEMENT TOOLS FOR MAINTAINING AN UPDATED REPOSITORY OF TOXICOLOGICALLY RELEVANT INFORMATION

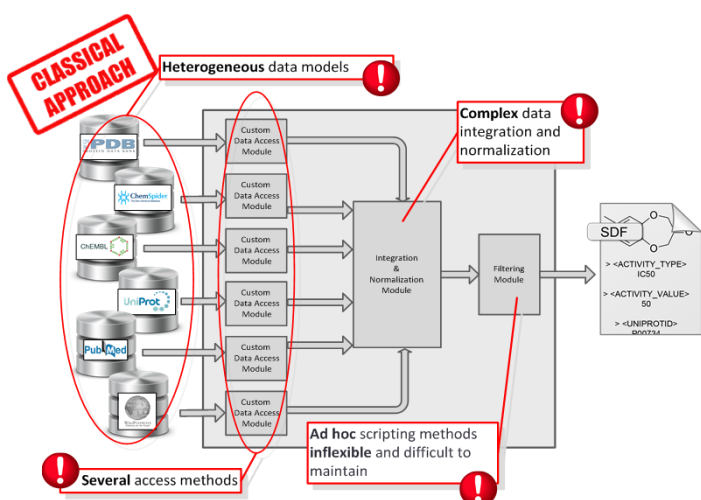Oriol L. Massaguer, Ferran Sanz, Manuel Pastor

Research Programme on *Biomedical Informatics* (GRIB), IMIM, Universitat Pompeu Fabra, Barcelona, Spain.

In this work, we used the Open PHACTS platform and framework to build a data crawling application which will distil automatically series of compounds annotated against toxicological endpoints from multiple public sources.

This information is automatically curated automatically using multi step ad hoc protocols and used to keep an updated local repository, in a format suitable for the building predictive models.

Obtaining bioactivity data for the building of in silico toxicity models

In the eTOX project [1,2] we face the necessity of obtaining bioactivity data for building in silico toxicity predictive models, typically series of compounds annotated with experimental bioactivity data about selected anti targets. The classical approach to compile manually and integrate the information from several data sources:

The Open PHACTS project [3,4,5] is building a platform providing access to the vast amount of information open resources in the areas of pharmaceutical research that can be found in the Internet.

The platform makes use of semantic web [6] technologies that allow a fully integrated access to all information without the need of align and integrate the data from a myriad of heterogeneous data sources.

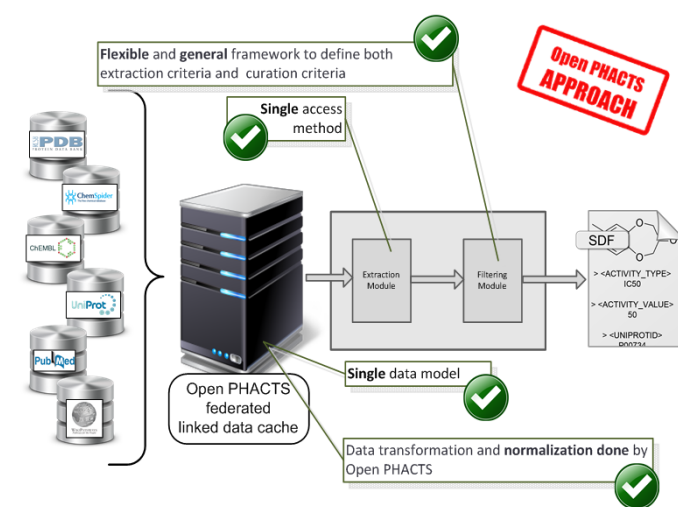With this approach the complexity of the system development is largely decreased.

Complex process to access an integrate data from several heterogeneous data sources.

Every time we want to extract new data the process have to repeat all the steps. Every time we want to add a new data source we have to modify the whole system.

The access to the data sources and normalization is responsibility of Open PHACTS.

The system is resilient to changes in data sources such as changes in access methods or addition of new data sources.

Solution proposed: data crawling and filtering tool based on Open PHACTS

The system developed allows the user to define data extraction jobs based on user-defined criteria such as target, activity type, activity values range or data sources

The user can define filtering protocols based on structural chemical criteria. The platform provides some prebuilt filters but the user can create new ones.

The user can export bioactivity and compound data to SDF or plain text of both raw or curated data.
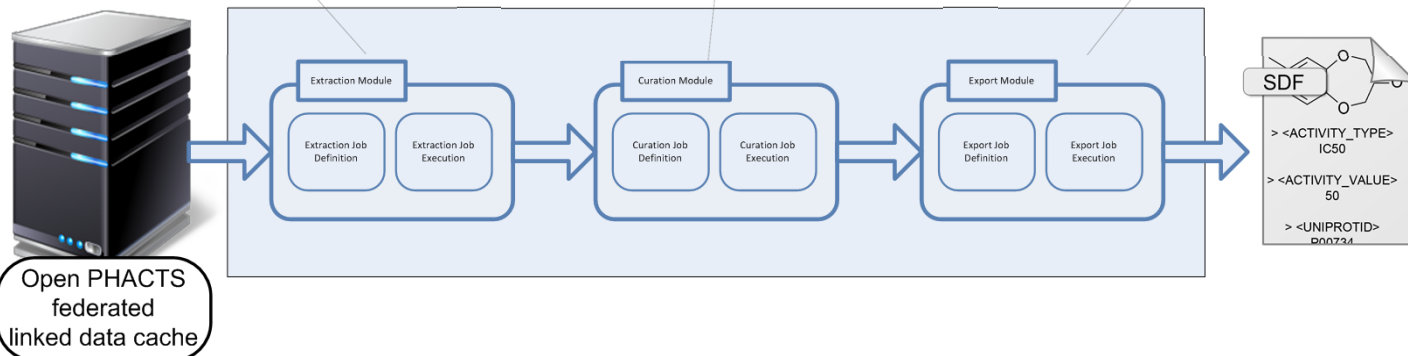
References

1. IMI-JU project: Integrating bioinformatics and chemoinformatics approaches for the development of expert systems allowing the *in silico* prediction of toxicities (eTOX). http://www.etoxproject.eu
2. Briggs K, Cases M, Heard DJ, Pastor M, Pognan F, Sanz F, Schwab CH, Steger-Hartmann T, Sutter A, Watson DK, Wichard JD. Inroads to Predict in Vivo Toxicology - An Introduction to the eTOX Troject. *Int J Mol Sci* 2012;**13**:3820-46.
3. Mons B et al. The Value of Data. Nat Genet 2011;**43**:281–283.
4. Blomberg N et al. Knowledge Driven Drug Discovery goes Semantic. *EFMC Yearbook* 2011.
5. Wild DJ, Ding Y, Sheth AP, Harland L, Gifford EM, Lajiness MS. Systems chemical biology and the Semantic Web: what they mean for the future of drug discovery research. *Drug Discov Today* 2011 May;**17**(9-10):469-474.
6. Berners-Lee T, Hendler J, Lassila O. The Semantic web. *SCIAM* 2001;**284**(5):34-43.

Acknowledgments