

Open PHACTS

Summary of Community Workshop 3

“Introducing Open PHACTS”

University of Vienna
30 August 2012

Content

| | |
|--|----|
| Introduction | 3 |
| Introducing Open PHACTS | 4 |
| The Open PHACTS infrastructure | 5 |
| The Open PHACTS Explorer | 6 |
| Open PHACTS Example Applications..... | 7 |
| PharmaTrek..... | 8 |
| Polypharmacology browser: GARField | 9 |
| Target dossier | 10 |
| ChemBio Navigator | 11 |
| Discussion & Feedback | 12 |
| How to get involved | 12 |

Introduction

The Open PHACTS consortium is building a freely available web platform to integrate pharmacological data from a variety of information resources, along with tools and services to question this integrated data. It will give researchers in industry and academia the ability to look up answers to complex research questions over a wide range of data sources, and provide a unique open framework to mix public and proprietary data. The project development is being directed by example questions sourced from expert researchers, and the open framework encourages innovative free and commercial applications to be built on the underlying platform. Open PHACTS will offer a production-grade API that commercial software providers can license and use to augment their own product. There are also opportunities for Associated Partners who want to do some more specific development work together with us, leading to a Development Partnership with the Open PHACTs project.

On August 30, 2012 the 3rd Open PHACTS Community Workshop "Introducing Open PHACTS" was held in Vienna, co-located with the 19th EuroQSAR conference, where the Open PHACTS consortium publicly presented the upcoming public Beta releases from the Open PHACTS project. The workshop introduced the technical and scientific approaches driving the project and demonstrated the web-based Explorer interface to the underlying Open PHACTS platform, built to answer specific research questions prioritised by the consortium. Additionally, four Open PHACTS exemplar applications that represent specialized interfaces were introduced. The example applications are being developed by consortium members, sit on top of the Open PHACTS platform and aim to address specific needs in the field of drug discovery.

The workshop was attended by 65 scientists, comprising representatives of software vendors (Inte:Ligand, OpenEye, ChemAxon, Tripos, Accelrys), participants of the EuroQSAR, as well as members of the Open PHACTS consortium. The Open PHACTS Explorer as well as 2 of the example applications were presented in a live demo. For in depth discussions, Open PHACTS was also present with an exhibition booth throughout the EuroQSAR conference.

Program:

- Introducing Open PHACTS (Bryn Williams-Jones)
- The Open PHACTS Infrastructure (Lee Harland)
- The Open PHACTS Explorer (Lee Harland/Paul Groth)
- Open PHACTS Example Applications
 - PharmaTrek (Jordi Mestres)
 - GARField (Louis Wich)
 - Target Dossier (Victor de la Torre)
 - ChemBio Navigator (Christian Lemmen)
- Discussion & Feedback



Picture 1: Participants of 3rd Open PHACTS Community Workshop

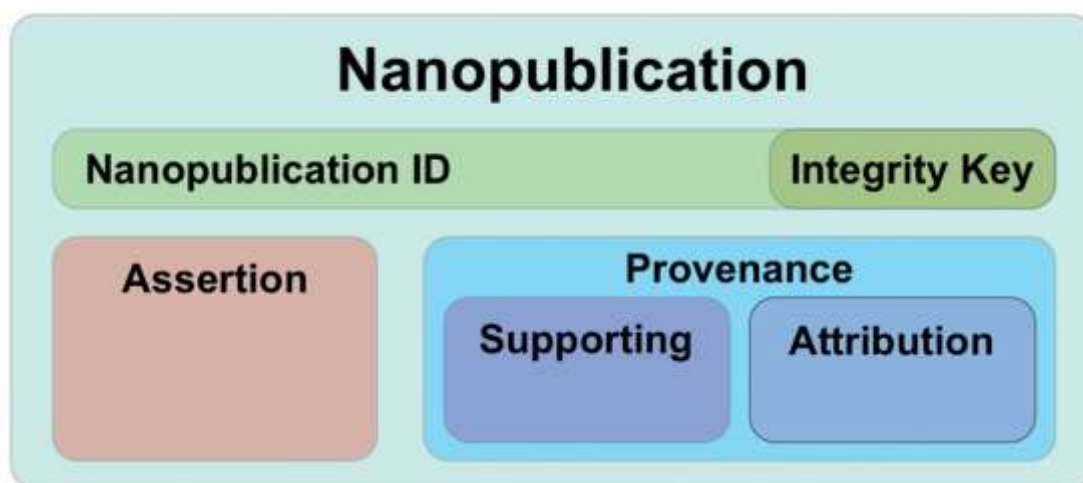
Introducing Open PHACTS

Open PHACTS acts like a search engine, allowing querying of data via a single platform while retaining provenance and traffic back to the original data source. Provenance is critical – users need to know where every data point comes from, and will visit the source. By promoting best practices for data publication and sharing, we want to simplify and clarify many of the problems around use and reuse of data from different sources.

Open PHACTS will be adopting a licensing framework which will be applicable to other, similar, data integration projects. This will provide clarity for the data sources and for the end user. The consortium wants to work with data providers to expose and enhance their data, and will build quality feedback mechanisms to help all its partners.

The power of nanopublication to capture core scientific assertions and promote data citeability will be demonstrated, and the project has already published nanopublication RDF guidelines. The first release of the Open PHACTS platform will include public sources of data and ontologies, but the licensing framework will allow the inclusion of different licenses, including the ability to include proprietary or commercial data. The project is building a window through which to query many data sources while retaining provenance.

Nanopublication Specifications



Picture 2: Nanopublication Specifications

Nanopublications provide support for provenance of data; credit to data providers and also allow user-annotations to be incorporated into the live system

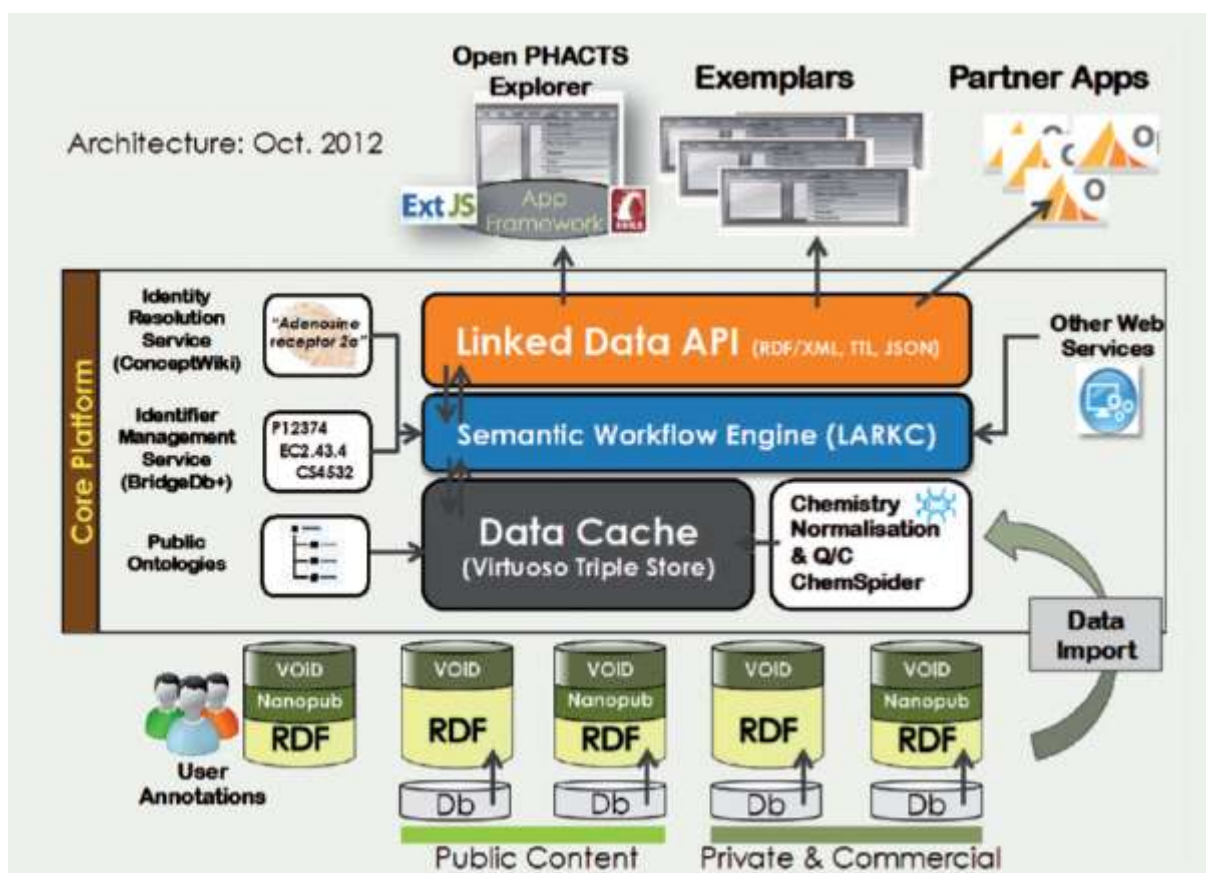
Timeline

| | |
|----------------------|--|
| November 2012 | Planned public release of Beta version 1.0 of Explorer |
| April 2013 | Alpha availability of the Open PHACTS API |
| 4Q 2013 | Beta version 2.0 of Explorer availability |
| 1Q 2014 | Final project release of Explorer |

The Open PHACTS infrastructure

The Open PHACTS platform uses semantic technologies to provide a robust, adaptable framework for integration of multiple data sources into one coherent API. While the project has a pharmacological focus, it will create a set of modular, reusable software components that can be used to address other scientific challenges. Open PHACTS is promoting and augmenting existing open standards and ontologies, and are demonstrating their use in a large scale, real world application. The Open PHACTS platform will be production software: data sources will be maintained and regularly updated, and the system is hosted by Open Link, a professional semantic data company.

Implementation architecture



Picture 3: Implementation architecture

Open PHACTS is markedly increasing the quality and flexibility of mapping between different identifiers in life science data – though developments on an Identity Mapping Service, the use of the ConceptWiki, and the publication of open data mappings.

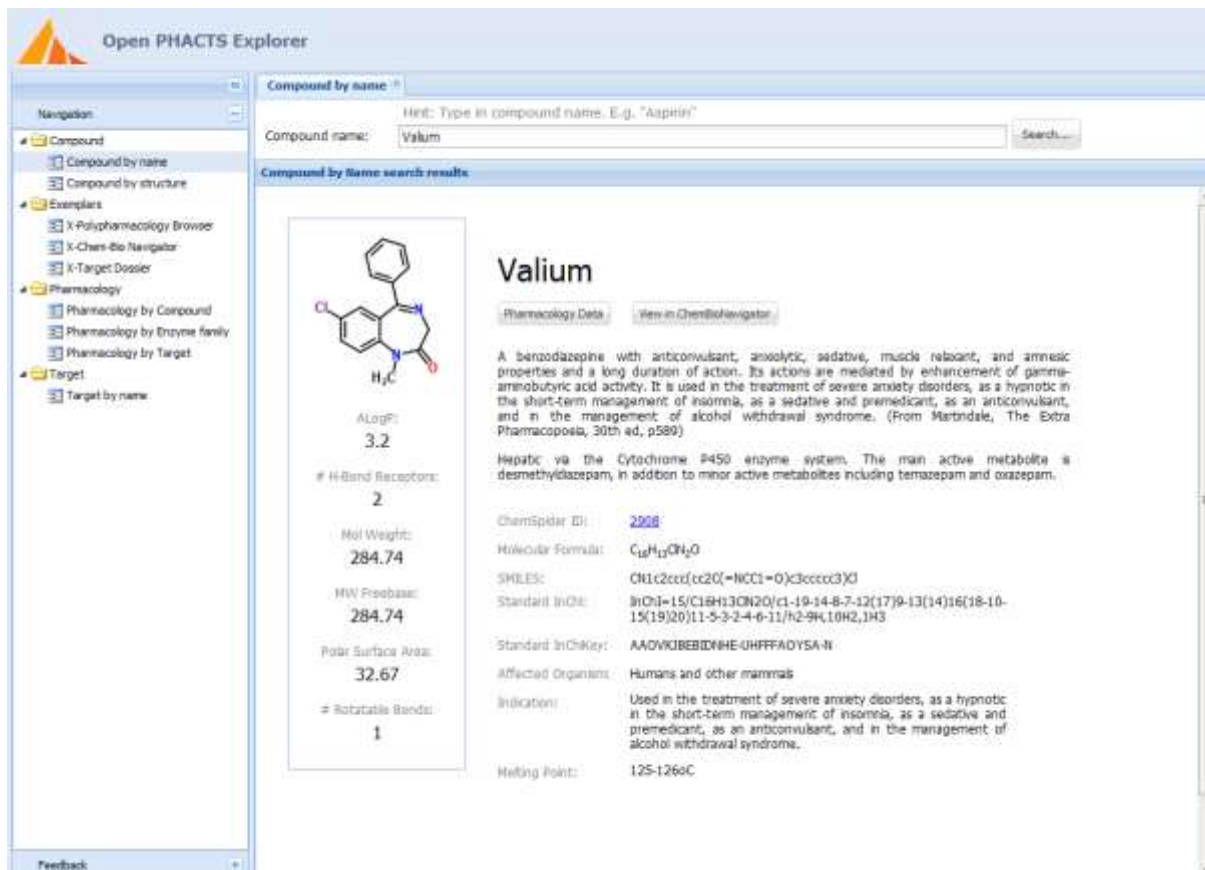
The Open PHACTS Explorer

The Open PHACTS Explorer provides a user interface to the Open PHACTS platform and is being built to answer the critical pharmacology questions defined by eight major pharmaceutical companies.

The first version of the platform includes data from ChEMBL, ChEBI, Uniprot, Gene Ontology, ChemSpider, WikiPathways, DrugBank, ENZYME, BridgeDB, predicted physical property data from ACD/Labs, and more.

The Explorer provides a way to query these up-to-date data sources; the integration process includes chemistry validation, and particularly deals with multiple identifiers for the same concept. The platform allows for rapid addition of new data sources, and the development of the platform and the Explorer has been use-case driven and tested by users in industry and academia.

It is being built, by experts, to offer a structured view of integrated pharmacological data. Open PHACTS gives you the ability to look up answers to complex research questions over a wide range of data sources. Unlike other tools it is built on an open vendor-neutral framework and allows license-compliant mixing of public and proprietary data with retained provenance.



The screenshot shows the Open PHACTS Explorer interface. On the left is a navigation pane with categories like Compound, Exemplars, Pharmacology, and Target. The main area displays search results for 'Valium'. It includes a chemical structure, key properties (AlogP: 3.2, # H-Bond Receptors: 2, Mol Weight: 284.74, MW Freebase: 284.74, Polar Surface Area: 32.67, # Rotatable Bonds: 1), and detailed pharmacology data such as ChemSpider ID (2908), Molecular Formula (C₁₆H₁₂ClN₂O), SMILES, Standard InChI, Standard InChIKey, Affected Organisms (Humans and other mammals), Indication, and Melting Point (125-126°C).

Picture 4: Open PHACTS Explorer

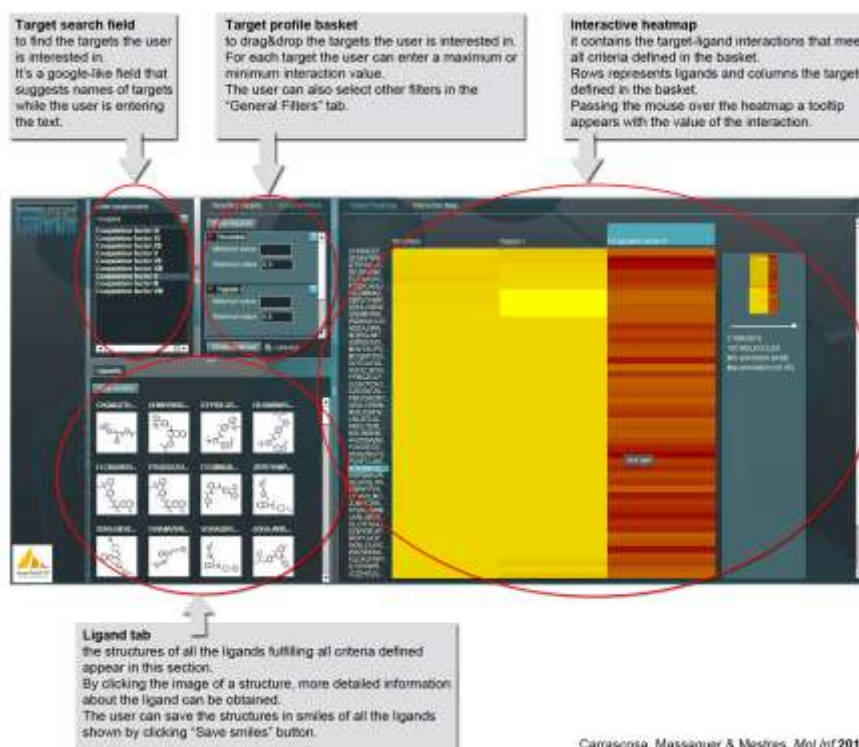
Open PHACTS Example Applications

Open PHACTS will offer a production-grade API that commercial software providers can license and use to augment their own product, generating revenue and long term sustainability. The API functions include general protein & compound information; pharmacology by target or compound; pharmacology by taxonomy, including ChEBI, GO, ENZYME and more.

Four example applications are being developed in the project to show how the data within the platform can be used to generate new insights:

PharmaTrek

PharmaTrek, developed by Consorci Mar Parc de Salut de Barcelona (PSMAR), proposes new mechanisms to navigate the Pharmacological space in a more interactive and flexible way. PharmaTrek is an integrative and interactive web application that will allow scientists to extract new knowledge from the Open Pharmacological Space created by OpenPHACTS. The main goal is to provide visual tools that allow the user to define custom questions, that is, that the users can create their own questions that will be answered in real time. The questions are related with the biological activity between drugs and targets.



Carrascosa, Massaquer & Mestres. Mol Inf 2012.

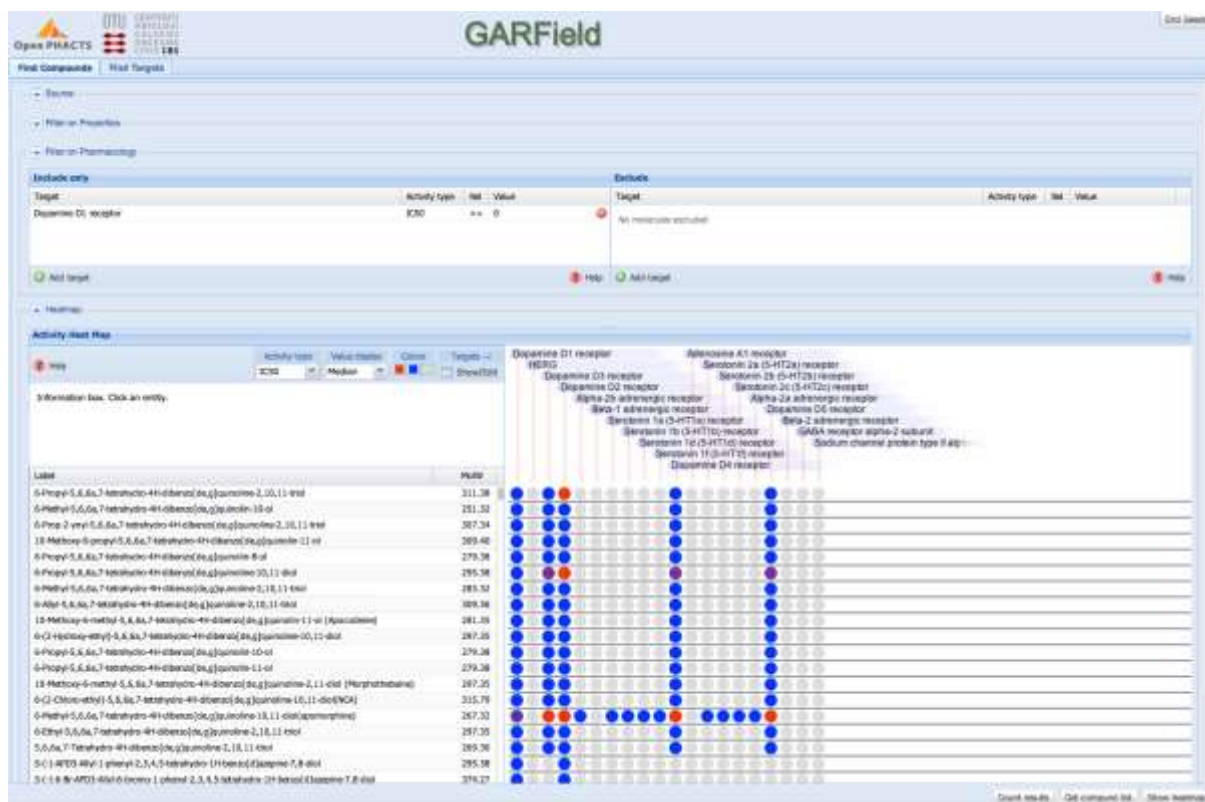
Picture 5: PharmaTrek



Picture 6: Jordi Mestres demonstrating PharmaTrek

Polypharmacology browser: GARField

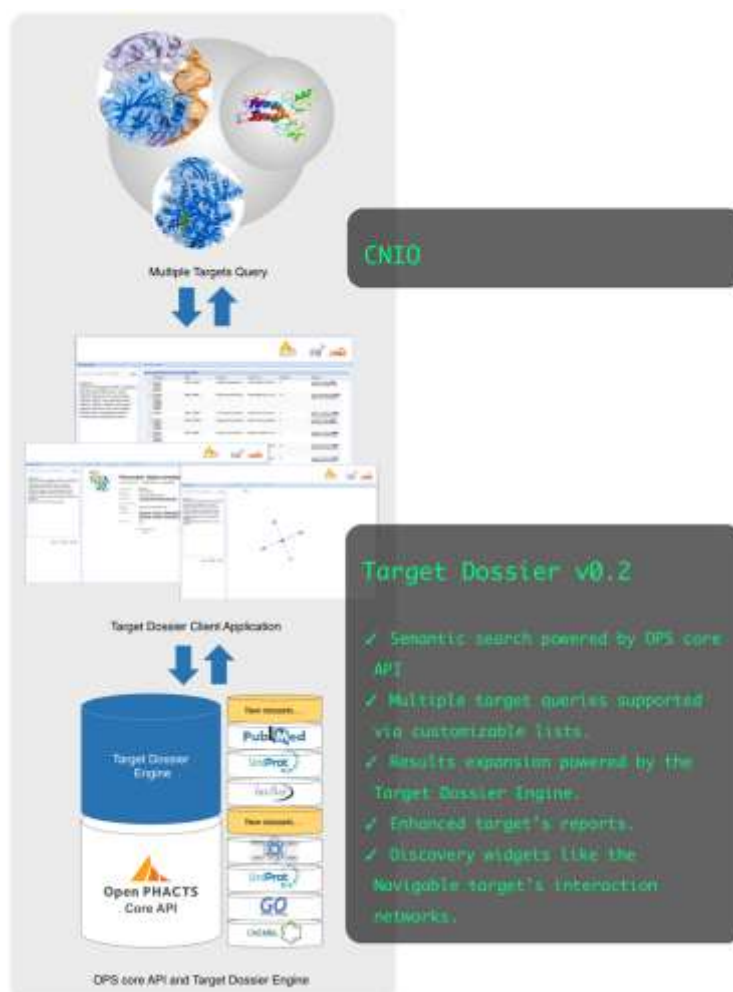
GARField, developed by the Technical University of Denmark, is a tool that facilitates the pharmacological profiling of small molecules and biological targets through the Open PHACTS services. It supports advanced searches for compounds that pass given criteria, e.g. fulfillment of certain chemical properties, and also filtering of the compounds by interaction with certain targets (for certain activity types). Likewise, the researcher can search for targets in similar queries, i.e. filtering by bioactivity to compounds. Results are presented visually in an interaction matrix. Besides the browsing capabilities, GARField will be an online platform for several bioactivity prediction algorithms, and so far implements the Similarity Ensemble Approach (Keiser et al. Nat. Biotechnol. 2007).



Picture 7: GARField

Target dossier

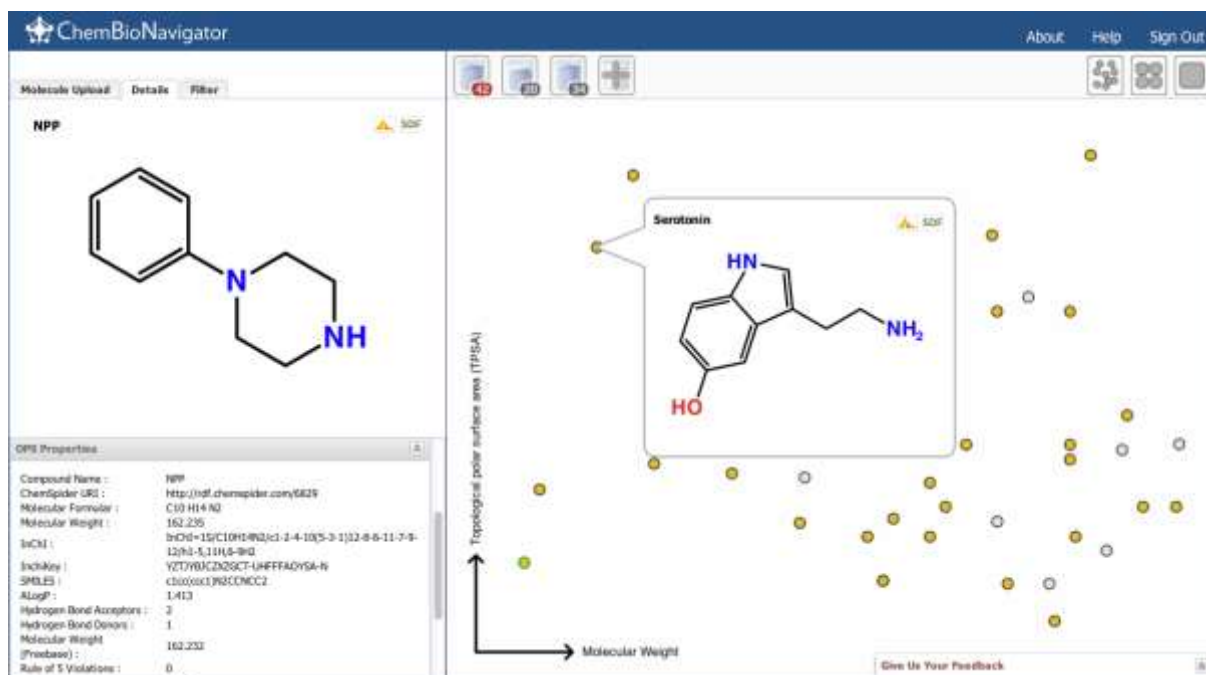
The Target Dossier, developed by the Spanish National Cancer Research Centre, uses the Open PHACTS platform for building comprehensive views of pharmacologically relevant targets to answer questions regarding druggability, tissue expression profiles and implications in pathways, disease states and physiological mechanisms. The Target Dossier will provide a decision support platform for target selection and progression.



Picture 8: Target Dossier

ChemBio Navigator

The ChemBioNavigator (CBN), developed by BioSolveIT, allows navigating at the interface of chemical and biological data. It is tailored for applications in pharmaceutical research. CBN lets you easily browse through sets of compounds. Different sorting and plotting options offer a quick and intuitive overview of the physio-chemical characteristics of the compounds. At the same time an in-depth analysis of individual molecules is supported using the versatile data available from the OPS platform. Hyperlinks into the original data sources as well as into the OPS Explorer allow for further investigation.



Picture 9: ChemBio Navigator

Discussion & Feedback

There was a discussion why the example applications look different from each other. This prompted an action item in the subsequent Steering Committee Meeting: Branding of Applications: How do we communicate clearly that Applications are developed by different organizations (under the umbrella of the project) and that they have their own identity/character (e.g. they don't use the same GUI design philosophy)?

How to get involved

Software vendors as well as other parties active in the area may join the Open PHACTS community by becoming an Associated Partner. Associated Partners will be the first to hear about the latest developments in the Open PHACTS project, and will also have the opportunity to present ideas and use cases to the core Open PHACTS team.

There are opportunities for Associated Partners who want to do some more specific development work together with the project team (e.g. develop APIs, new data, algorithms etc), leading to a Development Partnership with the Open PHACTS project.

For more information visit www.openphacts.org.