

Open PHACTS

OPS Deliverables

D 1.6. Deliver OPS domain vocabularies for version 1 and Phase 2 pilots

D 1.7. Deliver OPS Commons vocabularies for version 1 and Phase 2 pilots

Prepared by UNIMAN, UM, NBIC, SIB, CD
Approved by UNIMAN, UM, NBIC, SIB, CD, Pfizer, VUA

April 2012
Version 1.0

Project title: An open, integrated and sustainable chemistry, biology and pharmacology knowledge resource for drug discovery

Instrument: IMI JU

Contract no: 115191

Start date: 01 March 2011

Duration: 3 years

Nature of the Deliverable	
Report	x
Prototype	x
Other	
Dissemination level	
Public dissemination level	x
For internal use only	

Open PHACTS	Deliverables: Deliver OPS domain vocabularies for version 1 and Phase 2 pilots. Deliver OPS Commons vocabularies for version 1 and Phase 2 pilots.	Deliverable: D1.6 & D1.7	
IMI - 115191	Authors: Brenninkmeijer (UNIMAN), Chichester (SIB), Evelo (UM), Goble (UNIMAN), Gray (UNIMAN), Harland (CD), Waagmeester (UM), Willighagen (UM)	Version: 1.0	2 / 13

Definitions

- Partners of the Open PHACTS Consortium are referred to herein according to the following codes:

Pfizer – Pfizer limited – **Coordinator**

UNIVIE – Universität Wien – **Managing entity of IMI JU funding**

DTU – Technical University of Denmark – DTU

UHAM – University of Hamburg, Center for Bioinformatics

BIT – BioSolveIT GmbH

PSMAR – Consorci Mar Parc de Salut de Barcelona

LUMC – Leiden University Medical Centre

RSC – Royal Society of Chemistry

VUA – Vrije Universiteit Amsterdam

CNIO – Spanish National Cancer Research Centre

UNIMAN – University of Manchester

UM – University of Maastricht

ACK – ACKnowledge

USC – University of Santiago de Compostela

UBO – Rheinische Friedrich-Wilhelms-Universität Bonn

AZ – AstraZeneca

GSK – GlaxoSmithKline

Esteve – Laboratorios del Dr. Esteve, S.A.

Novartis – Novartis

ME – Merck Serono

HLU – H. Lundbeck A/S

E.Lilly – Eli Lilly

- Grant Agreement:** The agreement signed between the beneficiaries and the IMI JU for the undertaking of the Open PHACTS project.
- Project:** The sum of all activities carried out in the framework of the Grant Agreement.
- Work plan:** Schedule of tasks, deliverables, efforts, dates and responsibilities corresponding to the work to be carried, out as specified in the Grant Agreement.
- Consortium:** The Open PHACTS Consortium composed of the above-mentioned legal entities.
- Project Agreement:** Agreement concluded amongst Open PHACTS participants for the implementation of the Grant Agreement. Such an agreement shall not affect the parties' obligations to the Community and/or to one another arising from the Grant Agreement.

Open PHACTS	Deliverables: Deliver OPS domain vocabularies for version 1 and Phase 2 pilots. Deliver OPS Commons vocabularies for version 1 and Phase 2 pilots.	Deliverable: D1.6 & D1.7	
IMI - 115191	Authors: Brenninkmeijer (UNIMAN), Chichester (SIB), Evelo (UM), Goble (UNIMAN), Gray (UNIMAN), Harland (CD), Waagmeester (UM), Willighagen (UM)	Version: 1.0	3 / 13

1 General principles concerning vocabularies in Open PHACTS

This Open PHACTS deliverable provides an overview of the vocabularies that are currently being used by the system (Version 1, released March 2012) and those that are anticipated to be used in the future (Phase 2, release September 2012). This report may offer some general guidance as to the OPS preferred vocabularies but it should be understood that these are not being enforced as strict vocabulary standards in regards to data encoding in RDF. This document should be read in conjunction with the RDF publication guidelines [1].

The vocabularies currently in use are functioning as part of two different uses cases pertaining to WP1. The first is returning concepts that match a user entered search term in the interface. Here the concepts are entries in the vocabulary of discourse, which tie synonyms with a definition, and are being used for information retrieval. The second use case is the encoding of datasets in RDF using predicates from open, publicly available vocabularies and ontologies. Here the vocabulary terms describe the attributes of the data.

It should be noted that as the project evolves, this list of vocabularies for encoding RDF may change considerably as we learn more about the limitations of particular vocabularies but in all cases we will be guided by our principle of "not reinventing the wheel" and looking to use community vocabularies unless it is absolutely necessary to create our own. For the latest recommendations, the interested reader should consult the latest version of the RDF publication guidelines [1], and the wiki page

https://wiki.openphacts.org/index.php/Collection_of_URIs.

2 Version 1

Version 1 of the OPS system contains a limited number of data sources. These were published in RDF using a suitable domain and common vocabularies. The concepts covered by the datasets were captured to support their presentation in the user interfaces. Relationships between the data instances in different datasets were captured using a suitable vocabulary.

2.1 Domain Vocabularies

Domain vocabularies (shown in the table below) in Version 1 were used to select compounds and targets of interest in the user interface. Mapping between the concepts present in the different vocabularies was done to reduce redundancy.

Vocabulary	Contents
Medical Subject Headings (MeSH) [part of the Unified Medical Language System] http://purl.bioontology.org/ontology/MSH	Names of biological entities and processes, compounds

Open PHACTS	Deliverables: Deliver OPS domain vocabularies for version 1 and Phase 2 pilots. Deliver OPS Commons vocabularies for version 1 and Phase 2 pilots.	Deliverable: D1.6 & D1.7	
IMI - 115191	Authors: Brenninkmeijer (UNIMAN), Chichester (SIB), Evelo (UM), Goble (UNIMAN), Gray (UNIMAN), Harland (CD), Waagmeester (UM), Willighagen (UM)	Version: 1.0	4 / 13

GeneOntology (GO) [part of the Unified Medical Language System] http://purl.org/obo/owl/GO#	Biological processes, molecular functions, and cellular components
SwissProt www.uniprot.org	Names and synonyms of proteins
ChemSpider www.chemspider.com	Names and synonyms of compounds

The datasets used in Version 1 have been published by multiple sources using different namespaces for the URIs. The table below gives the namespaces that have been used.

Dataset	Namespace
ChEMBL	http://chem2bio2rdf.org/chembl/resource
ChemSpider	http://rdf.chemspider.com/
ConceptWiki	http://www.conceptwiki.org/concept/
DrugBank	http://www4.wiwiss.fu-berlin.de/drugbank/resource/drugbank/
ENZYME	http://purl.uniprot.org/enzyme/

2.2 Commons Vocabularies

The table below details the commons vocabularies used to describe the data sets and the linksets that related them.

Vocabulary	Description	Usage
DBPedia http://dbpedia.org/ontology/	RDF version of concepts described in Wikipedia.	DrugBank
Dublin Core (DC) http://dublincore.org/	Vocabulary of metadata terms to describe the authorship of a document.	ChemSpider, Linksets
Friend of a Friend (FOAF) http://www.foaf-project.org/	Vocabulary to describe people.	ChEMBL, ChemSpider, DrugBank
Resource Description Framework (RDF) http://www.w3.org/1999/02/22-rdf-syntax-ns#	Data representation system	Brenda, ChEMBL, ChemSpider, DrugBank

Open PHACTS	Deliverables: Deliver OPS domain vocabularies for version 1 and Phase 2 pilots. Deliver OPS Commons vocabularies for version 1 and Phase 2 pilots.	Deliverable: D1.6 & D1.7	
IMI - 115191	Authors: Brenninkmeijer (UNIMAN), Chichester (SIB), Evelo (UM), Goble (UNIMAN), Gray (UNIMAN), Harland (CD), Waagmeester (UM), Willighagen (UM)	Version: 1.0	5 / 13

Resource Description Framework Schema (RDFS) http://www.w3.org/2000/01/rdf-schema#	Provides basic predicates for sub-classing.	Brenda, ChEMBL, ChemSpider, DrugBank
Simple Knowledge Organisation System (SKOS) http://www.w3.org/TR/skos-reference/	Vocabulary for relating concepts across data sources.	Linksets
Vocabulary of Interlinked Datasets (VoID) http://rdfs.org/ns/void#	Vocabulary to describe a dataset.	ChemSpider, Linksets
Web Ontology Language (OWL) http://www.w3.org/2002/07/owl#	Advanced knowledge representation.	ChEMBL, DrugBank
XML Schema (XSD) http://www.w3.org/2001/XMLSchema#	Provides basic types, e.g. integer	ChEMBL, DrugBank

3 Phase 2 Pilot

3.1 Scientific and Domain Vocabularies

The BioPortal website, <http://bioportal.bioontology.org/>, maintained by the NCBO, an Open PHACTS associated partner, hosts ontologies that have been published by the community (293 in total). As not all of these ontologies are commonly used, and often alternatives exist for the same domain, an exhaustive overview of ontologies is not made. Instead, we here list a small selection of those that may be of interest to the Open PHACTS project:

- Biological pathways
 - BioPax (BP): formal representation of biological pathways.
<http://www.biopax.org/>
 - Pathway ontology: captures biological networks and the relationships between them ftp://rgd.mcw.edu/pub/data_release/ontology_obo_files/
- Assays
 - BioAssay Ontology (BAO): describes screening assays including high-throughput screening <http://www.bioassayontology.org/>
 - Ontology for Drug Discovery Investigations: describes drug discovery investigations, including assays <http://purl.org/ddi/home>
- General Biology
 - Gene Ontology (GO), <http://www.geneontology.org/>
 - PSI-MOD (MOD), <http://www.psidev.info/>
 - Cell line ontology, Cellosaurus (CVCL), ftp://ftp.nextprot.org/pub/current_release/controlled_vocabularies/cellosaurus.txt

Open PHACTS	Deliverables: Deliver OPS domain vocabularies for version 1 and Phase 2 pilots. Deliver OPS Commons vocabularies for version 1 and Phase 2 pilots.	Deliverable: D1.6 & D1.7	
IMI - 115191	Authors: Brenninkmeijer (UNIMAN), Chichester (SIB), Evelo (UM), Goble (UNIMAN), Gray (UNIMAN), Harland (CD), Waagmeester (UM), Willighagen (UM)	Version: 1.0	6 / 13

- Protein Ontology <http://bioportal.bioontology.org/ontologies/1052>
- Medicine & Clinical
 - Translational Medicine Ontology (TMO) <http://esw.w3.org/HCLSIG/PharmaOntology>
 - Drug Ontology Project for Elsevier (DOPE) <http://www.w3.org/2001/sw/sweo/public/UseCases/Elsevier/>
 - International Classification of diseases (ICD-10) <http://www.who.int/classifications/icd/en/>
 - Systematized Nomenclature of Medicine Clinical Terms (SNOMED) <http://www.connectingforhealth.nhs.uk/systemsandservices/data/uktc/snomed>
 - Medical Dictionary for Regulatory Activities (MedDRA) <http://www.meddransso.com/>
 - Human Disease Ontology (HDO) <http://bioportal.bioontology.org/ontologies/1009>
 - Disease Ontology <http://disease-ontology.org>
 - Ontology of Clinical Research (OCRe) <http://bioportal.bioontology.org/ontologies/1076>
- Chemical compounds
 - Chemical Entities of Biological Interest (ChEBI): classification of compounds by functionality and functional groups <http://www.ebi.ac.uk/ontology-lookup/browse.do?ontName=ChEBI>
 - Chemical information ontology (CHEMINF): representation of chemical structures and properties of chemicals <http://code.google.com/p/semanticchemistry/>
 - ChemAxiom – An Ontological Framework for Chemistry in Science <http://bit.ly/axVpjp>
 - Name reaction ontology (RXNO), Chemical methods ontology (CMO), Molecular processes ontology (MOP) <http://www.rsc.org/ontologies/>

Recommendations for OPS domain vocabularies to be used in the phase 2 pilot are given in the table below:

Vocabulary	Usage	Justification
BioPax: http://www.biopax.org/	Biological pathways and protein-protein interactions: Wikipathways, ihop nanopublications	Eleven databases include BioPax in their export
BioAssay ontology: http://www.bioassayontology.org/	High-throughput screening assays	Is being used by EPFIA to annotate their HTS
Chemical entities of biological interest (ChEBI), http://www.ebi.ac.uk/chebi/	Dictionary of molecular entities focused on 'small' chemical compounds	A structured classification of chemical compounds of

Open PHACTS	Deliverables: Deliver OPS domain vocabularies for version 1 and Phase 2 pilots. Deliver OPS Commons vocabularies for version 1 and Phase 2 pilots.	Deliverable: D1.6 & D1.7	
IMI - 115191	Authors: Brenninkmeijer (UNIMAN), Chichester (SIB), Evelo (UM), Goble (UNIMAN), Gray (UNIMAN), Harland (CD), Waagmeester (UM), Willighagen (UM)	Version: 1.0	7 / 13

		biological relevance that is endorsed by many partners
GeneOntology, http://www.geneontology.org/	Describes gene product characteristics	Most widely used ontology in biology
NCBI Organism Taxonomy http://bioportal.bioontology.org/ontologies/1132	Organism heirarchy	Most widely used organism classification, already used in our data sources
CHEMINF http://code.google.com/p/semanticchemistry/	Cheminformatics Ontology	Used by partners to represent chemical entity descriptors and identifiers

Some specific vocabularies are listed below:

- General Science
 - Semanticscience Integrated Ontology (SIO) <http://code.google.com/p/semanticscience/wiki/SIO>
 - Information Artifact Ontology (IAO) <http://code.google.com/p/information-artifact-ontology/>
- Discourse:
 - Linked Science ontologies, <http://linkedscience.org/lsc/ns/>
 - Scientific Discourse Ontology (SWAN) <http://www.w3.org/TR/hcls-swan/>
 - Semantically-Interlinked Online Communities (SOIC) <http://sioc-project.org/>
- Experiments:
 - Investigations Studies Assays (ISA): describes linkages between experiments <http://isatab.sourceforge.net/>
 - EXPO: Ontology of Scientific Experiments <http://expo.sourceforge.net/>
 - Experimental Factor Ontology: describes transcript omics experiments, but being generalised to all experiments <http://www.ebi.ac.uk/efo>
 - Semantic science Integrated Ontology (SIO) <http://semanticscience.org/ontology/sio.owl>
 - Ontology for Biomedical Investigations (OBI): describes experimental conditions <http://obi-ontology.org/>
- Units
 - Units of Measure <http://bioportal.bioontology.org/ontologies/47017>
 - Units Ontology <http://bioportal.bioontology.org/ontologies/1650>
 - QUDT an ontology of scientific units <http://quidt.org>

Open PHACTS	Deliverables: Deliver OPS domain vocabularies for version 1 and Phase 2 pilots. Deliver OPS Commons vocabularies for version 1 and Phase 2 pilots.	Deliverable: D1.6 & D1.7	
IMI - 115191	Authors: Brenninkmeijer (UNIMAN), Chichester (SIB), Evelo (UM), Goble (UNIMAN), Gray (UNIMAN), Harland (CD), Waagmeester (UM), Willighagen (UM)	Version: 1.0	8 / 13

Recommendations for OPS scientific vocabularies to be used in the phase 2 pilot are given in the table below:

Name	Covers	Justification
SWAN: http://www.w3.org/TR/hcls-swan/	Discourse	Widely used
EFO: http://www.ebi.ac.uk/efo	Experiments	Focused ontology
QUDT: http://quidt.org	Scientific units	Extremely well thought out, much attention to detail, extensive coverage of common and not so common units

3.2 Commons Vocabularies

The OPS Platform itself needs vocabularies. The Open PHACTS consortium has worked on guidelines for publishing datasets as RDF [1] and for publishing nanopublications [2]. We aim to reuse existing standards and have involved ourselves in standards activities, e.g. the W3C Provenance Working Group¹. The vocabularies identified in the first stage of the project include:

- Minimum information:
 - Dublin Core (DC): Core attribution vocabulary
<http://dublincore.org/>
 - Minimum information model (MIMV): A minimal model for describing objects against a specification of Must/Should/Optional requirements
<http://purl.org/net/mim/ns>
 - Vocabulary of Interlinked Datasets (VOID) for describing datasets and the linksets that relate them
<http://rdfs.org/ns/void#>
- Provenance:
 - Dublin Core (DC): Core attribution vocabulary
<http://dublincore.org/>
 - Provenance Authoring and Versioning (PAV) ontology: Provides more fine-grained attribution than dublin core. Provide as part of the Semantic Web Applications in Neuromedicine (SWAN) project
<http://code.google.com/p/pav-ontology/>
 - Vocabulary of Interlinked Datasets (VOID) for describing datasets and the linksets that relate them
<http://rdfs.org/ns/void#>
 - Open Provenance Model (OPM):
<http://openprovenance.org/>

¹ http://www.w3.org/2011/prov/wiki/Main_Page accessed April 2012.

Open PHACTS	Deliverables: Deliver OPS domain vocabularies for version 1 and Phase 2 pilots. Deliver OPS Commons vocabularies for version 1 and Phase 2 pilots.	Deliverable: D1.6 & D1.7	
IMI - 115191	Authors: Brenninkmeijer (UNIMAN), Chichester (SIB), Evelo (UM), Goble (UNIMAN), Gray (UNIMAN), Harland (CD), Waagmeester (UM), Willighagen (UM)	Version: 1.0	9 / 13

- W3C's provenance ontology (PROV-O): process oriented provenance ontology
<http://www.w3.org/TR/prov-o/>
- Versioning:
 - Memento: technical framework that adds a time dimension to the HTTP protocol
<http://www.mementoweb.org/>
 - EvoOnt: software evolution ontology
<https://files.ifi.uzh.ch/ddis/oldweb/ddis/research/evoont/index.html>
 - Ontology Metadata Vocabulary (OMV) Change ontology extension
<http://omv2.sourceforge.net/>
 - SWAN: Provenance, authoring, and versioning ontology
<http://swan.mindinformatics.org/>
- Community:
 - Friend of a Friend (FOAF): describe social networks.
<http://xmlns.com/foaf/0.1/>
 - Semantically Interlinked Online Communities (SIOC) for linking discussion methods such as blogs, forums, and mailing lists
<http://sioc-project.org/>
 - OpenSocial: framework for developing social web applications
<http://www.opensocial.org/>
 - VIVO <http://vivoweb.org/download>
- Aggregation:
 - Open Archives Initiative Object Reuse and Exchange (OAI-ORE): enables the creation of digital objects that aggregate web resources.
<http://www.openarchives.org/ore/>
- Harvesting:
 - Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH):
<http://www.openarchives.org/pmh/>
- Citation:
 - Citation Typing Ontology (CITO) part of the Semantic Publishing and Referencing (SPAR) framework. Describes how different resources cite each other
<http://purl.org/spar/cito/>
 - Bibliographic Ontology (BIBO): represents bibliographic data, such as article and book references.
<http://purl.org/ontology/bibo/>
- Annotation:
 - Annotation Ontology (AO): adding comments, tags, notes to documents
<https://code.google.com/p/annotation-ontology/>
 - Open Annotation Collaboration (OAC): adding notes, commentary, and classifying documents
<http://www.openannotation.org/>
 - Open Annotation Community Group: reconciliation of AO and OAC.
<http://www.w3.org/community/openannotation/>
- Identifier:
 - Digital Object Identifiers (DOI) for data and documents
<http://www.doi.org/>

Open PHACTS	Deliverables: Deliver OPS domain vocabularies for version 1 and Phase 2 pilots. Deliver OPS Commons vocabularies for version 1 and Phase 2 pilots.	Deliverable: D1.6 & D1.7	
IMI - 115191	Authors: Brenninkmeijer (UNIMAN), Chichester (SIB), Evelo (UM), Goble (UNIMAN), Gray (UNIMAN), Harland (CD), Waagmeester (UM), Willighagen (UM)	Version: 1.0	10 / 13

- Open Researcher and Contributor ID (ORCID) for researchers.
<http://orcid.org/>
- Mappings:
 - Simple Knowledge Organisation System (SKOS): provides generic mapping relationships for information retrieval tasks
<http://www.w3.org/TR/skos-reference/>
 - Similarity ontology: proposed by Halpin *et al* (2010) as a more fine grained generic mapping ontology than SKOS.
- Knowledge representation
 - Resource Description Framework (RDF)
<http://www.w3.org/1999/02/22-rdf-syntax-ns#>
 - Resource Description Framework Schema (RDF-S)
<http://www.w3.org/2000/01/rdf-schema#>
 - Web Ontology Language (OWL)
<http://www.w3.org/2002/07/owl#>
 - Open Biological and Biomedical Ontologies (OBO)
<http://obofoundry.org/>

Versioning of ontologies is a difficult issue that has received much attention in the research literature. The Open PHACTS project does not aim to create or maintain ontologies, but to reuse existing ontologies. As such, ontology versioning is not considered.

Recommendations for OPS commons vocabularies to be used in the phase 2 pilot are given in the table below:

Name	Covers	Justification
W3C Open Annotation Community http://www.w3.org/community/openannotation/	Annotations	Merger of the two widely used annotation ontologies
Dublin Core http://dublincore.org/	Core attribution	Widely used vocabulary for attribution
Provenance Ontology http://www.w3.org/TR/prov-o/	Provenance	Covers a wide range of provenance use cases, expected to be widely adopted
Nanopublication Ontology http://www.nanopub.org/nschema	Nanopublication concepts	Only vocabulary that describes nanopublications
Data Cube http://publishing-statistical-data.googlecode.com/svn/trunk/specs/src/main/html/cube.html	Statistics	Representation of the ISO standard for exchanging statistical data

Open PHACTS	Deliverables: Deliver OPS domain vocabularies for version 1 and Phase 2 pilots. Deliver OPS Commons vocabularies for version 1 and Phase 2 pilots.	Deliverable: D1.6 & D1.7	
IMI - 115191	Authors: Brenninkmeijer (UNIMAN), Chichester (SIB), Evelo (UM), Goble (UNIMAN), Gray (UNIMAN), Harland (CD), Waagmeester (UM), Willighagen (UM)	Version: 1.0	11 / 13

Creative Commons Vocabulary http://wiki.creativecommons.org/CC_REL	Licensing	Licenses aimed at sharing open data
FOAF http://xmlns.com/foaf/0.1/	Social networks	Widely used vocabulary for describing interactions between individuals
DOI http://www.doi.org/	Document identifiers	Standard identifier scheme used by publishers for documents
ORCID http://orcid.org/	Author identifiers	Open author/contributor identifier scheme that identifies an individual rather than an individual at a specific institution
SKOS http://www.w3.org/TR/skos-reference/	Generic mappings	Captures the information retrieval semantics required for mapping concepts for search
Void http://rdfs.org/ns/void#	Dataset and linksets specifications	Captures many of the notions required to describe a dataset, its attributions, and its content
RDF http://www.w3.org/1999/02/22-rdf-syntax-ns#	Data representation system	Required for representing RDF data upon which everything builds
RDFS http://www.w3.org/2000/01/rdf-schema#	Basic knowledge representation	Widely used vocabulary with basic relationships
Web Ontology Language OWL http://www.w3.org/2002/07/owl#	Advanced knowledge representation	Provides the mechanism for defining formal relationships between concepts. Widely adopted.

<p style="text-align: center;">Open PHACTS</p>	<p>Deliverables: Deliver OPS domain vocabularies for version 1 and Phase 2 pilots. Deliver OPS Commons vocabularies for version 1 and Phase 2 pilots.</p>	<p>Deliverable: D1.6 & D1.7</p>	
<p>IMI - 115191</p>	<p>Authors: Brenninkmeijer (UNIMAN), Chichester (SIB), Evelo (UM), Goble (UNIMAN), Gray (UNIMAN), Harland (CD), Waagmeester (UM), Willighagen (UM)</p>	<p>Version: 1.0</p>	<p>12 / 13</p>

3.3 Example Usage

Wikipathways is a biological pathways dataset that will be made available as RDF for the September release. It uses the following recommended ontologies:

- Bibliography Ontology is used to describe bibliographic data, such as article and book references. <http://purl.org/ontology/bibo/>
<http://rdfs.org/ns/void#>
- The Citation Typing Ontology is used to describe how different resources cite each other. <http://purl.org/spar/cito/>
- Dublin Core is used to provide attribution metadata. <http://purl.org/dc/elements/1.1/>
- RDF Core vocabulary defined the foundation of the RDF. <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
- RDFS is an extension of RDF to introduce basic concepts allowing knowledge representation. <http://www.w3.org/2000/01/rdf-schema#>
- The Web Ontology Language extends RDFS with more formal concepts, at various levels, so that formal ontologies can be represented in RDF. <http://www.w3.org/2002/07/owl#>
- Friend-of-a-Friend is an ontology to describe social networks. <http://xmlns.com/foaf/0.1/>
- BioPax is an ontology to describe pathway data: <http://www.biopax.org/release/biopax-level3.owl>

ChEMBL is a chemical database containing bioactive molecules with drug-like properties. The previous conversion to RDF provided by the Chem2Bio2RDF project [3] has certain modelling short comings and has not been upgraded since version 8 of ChEMBL was released. A new conversion of ChEMBL based on version 13 will be used for the September release that uses the following recommended ontologies.

- Bibliography Ontology is used to describe bibliographic data, such as article and book references. <http://purl.org/ontology/bibo/>
- The Citation Typing Ontology is used to describe how different resources cite each other. <http://purl.org/spar/cito/>
- Dublin Core is used to provide attribution metadata. <http://purl.org/dc/elements/1.1/>
- RDF Core vocabulary defined the foundation of the RDF. <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
- RDFS is an extension of RDF to introduce basic concepts allowing knowledge representation. <http://www.w3.org/2000/01/rdf-schema#>
- The Web Ontology Language extends RDFS with more formal concepts, at various levels, so that formal ontologies can be represented in RDF. <http://www.w3.org/2002/07/owl#>
- The SemanticScience Integrated Ontology (SIO) and Chemical Information (CHEMINF) ontology for cheminformatics information.

Open PHACTS	Deliverables: Deliver OPS domain vocabularies for version 1 and Phase 2 pilots. Deliver OPS Commons vocabularies for version 1 and Phase 2 pilots.	Deliverable: D1.6 & D1.7	
IMI - 115191	Authors: Brenninkmeijer (UNIMAN), Chichester (SIB), Evelo (UM), Goble (UNIMAN), Gray (UNIMAN), Harland (CD), Waagmeester (UM), Willighagen (UM)	Version: 1.0	13 / 13

<http://code.google.com/p/semanticsscience/wiki/SIO>,

<http://bioportal.bioontology.org/ontologies/1444>

- Protein Ontology is used for protein targets. <http://pir.georgetown.edu/pro/>

4 References

- [1] Carina Haupt, Andra Waagmeester, Marc Zimmermann, and Egon Willighagen. Open PHACTS RDF Guidelines. Version 0.1, 29 February 2012.
<https://docs.google.com/document/d/1FPLK-Y1RbxpsiL3IHWq3ZN99pbu2E8V3Ms34IcGMxII/edit>
- [2] Erik Schultes, Christine Chichester, Kees Burger, Spyros Kotoulas, Antonis Loizou, Valery Tkachenko, Andra Waagmeester, Sune Askjaer, Steve Pettifer, Lee Harland, Carina Haupt, Colin Batchelor, Miguel Vazquez, José María Fernández, Jahn Saito, Andrew Gibson, and Louis Wich. The Open PHACTS Nanopublications Guidelines. Version 1.8, 14 February 2012.
- [3] B. Chen, X. Dong, D. Jiao, H. Wang, Q. Zhu, Y. Ding, and D.J. Wild. Chem2Bio2RDF: a semantic framework for linking and data mining chemogenomic and systems chemical biology data. BMC Bioinformatics, 11(1):255, 2010.