(towards the) Semantic annotation of the Laboratory Chemical Safety Summary in PubChem

Gang Fu[1], Jian Zhang[1], **Evan Bolton**[1], Jeremy Frey[2], Stuart Chalk[3], Mark Borkum[4], Leah McEwen[5]

1 **National Center for Biotechnology Information, Bethesda, MD, USA**

2 University of Southampton, Southampton, UK

3 Department of Chemistry, University of Florida, Jacksonville, FL, USA

4 Environmental Molecule Sciences Laboratory, PNNL, Richland, WA, USA

5 Clark Library, Cornell University, Ithaca, NY, USA

PubChem

PubChem Presentations

Monday, August 22

CINF 47: Practical issues in chemistry data sharing in PubChem
Room 112A – Convention Center, 10:50 am – 11:00 am

CINF 58: Chemistry data pain points: distilled, analyzed, and next steps
Room 112A – Convention Center, 1:55 pm – 4:10 pm

Tuesday, August 23

CINF 76: Open chemical information: Where now and how?
Room 112A/B – Convention Center, 4:25 pm – 4:50 pm

Wednesday, August 24

CINF 77: Users roundtable: Laboratory use cases for chemical safety information
Room 112A – Convention Center, 8:30 am – 8:45 am

CINF 80: Chemical safety and hazard information in PubChem
Room 112A – Convention Center, 9:35 am – 10:00 am

CINF 81: Semantic annotation of the laboratory chemical safety summary in PubChem
Room 112A – Convention Center, 10:15 am – 10:40 am

Thursday, August 25

CINF 93: Strategies to improve PubChem data quality and search effectiveness through data analysis
Room 112A – Convention Center, 9:15 am – 9:40 am

CINF 95: Hybrid search engine for chemical information in PubChem
Room 112A – Convention Center, 10:20 am – 10:45 am

OUTLINE

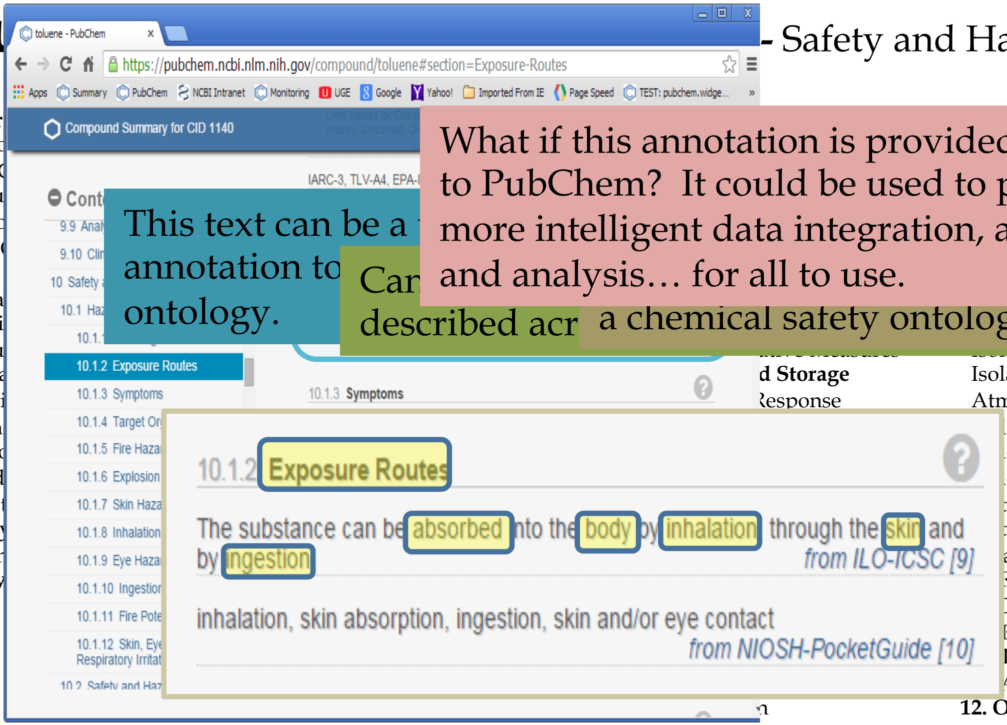
- ➢ PubChemRDF Overview
- ➢ Semantic Annotation of Physical Properties
- ➢ Semantic Annotation of Global Harmonized System
- ➢ PubChemRDF Use Cases
- ➢ Community involvement

How can PubChem help?

Well .. we have lots of data

Eureka!!!
Let’s Make a connected graph of knowledge

Pu - Safety and Hazards



Compound Summary for CID 1140

IARC-3, TLV-A4, EPA-I

10.1.3 Symptoms

10.1.2 Exposure Routes

The substance can be absorbed into the body by inhalation through the skin and by ingestion

from ILO-ICSC [9]

inhalation, skin absorption, ingestion, skin and/or eye contact

from NIOSH-PocketGuide [10]

This text can be a annotation to ontology.

Can described acr

mostly using a chemical safety ontology?

What if this annotation is provided back to PubChem? It could be used to power more intelligent data integration, access and analysis… for all to use.

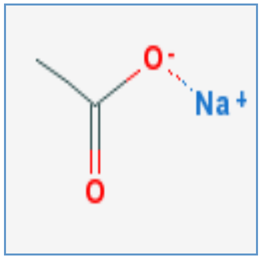
1. Hazar
Carcinc
Safety C
Exposu
Sympto
Target C
Cancer
Fire Ha
Explosi
Exposu
Skin Ha
Inhalati
Eye Ha
Ingestic
Hazard
Fire Po
Skin, Ey
Irritation
2. Safety
LEL
UEL
IDLH
REL
PEL
PEL-TWA
PEL-STEL
PEL-C
REL-TWA
REL-STEL
REL-C
Conversion
Flammability

3. First Aid Measures
First Aid
Fire First Aid
Explosion First Aid
Exposure First Aid
Inhalation First Aid
Skin First Aid
Eye First Aid

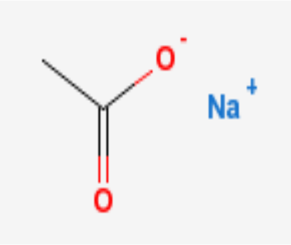
d Storage
Response
n
Ingestion Prevention
Protective Equipment and Clothing
8. Stability and Reactivity
Reactivities and Incompatibilities
9. Disposal Considerations
10.Transport Information
DOT Emergency Guidelines
Shipment Methods and

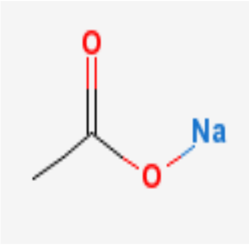
Isolation Distance
Atmospheric Standards
Standards
eral Drinking Water Standards
eral Drinking Water Guidelines
e Drinking Water Standards
e Drinking Water Guidelines
an Water Act Requirements
RCLA Reportable Quantities
CA Requirements
RA Requirements
RA Requirements
A Requirements
12. Other Safety Information
Safety References
Safety Notes
Toxic Combustion Products
Other Hazardous Reactions
Material Safety Data Sheet

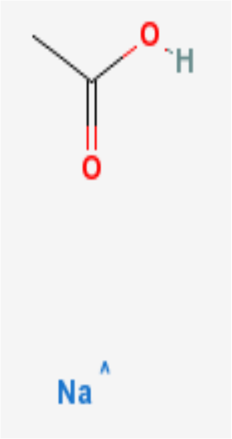
A chemical structure may be represented in many different ways



Sodium Acetate

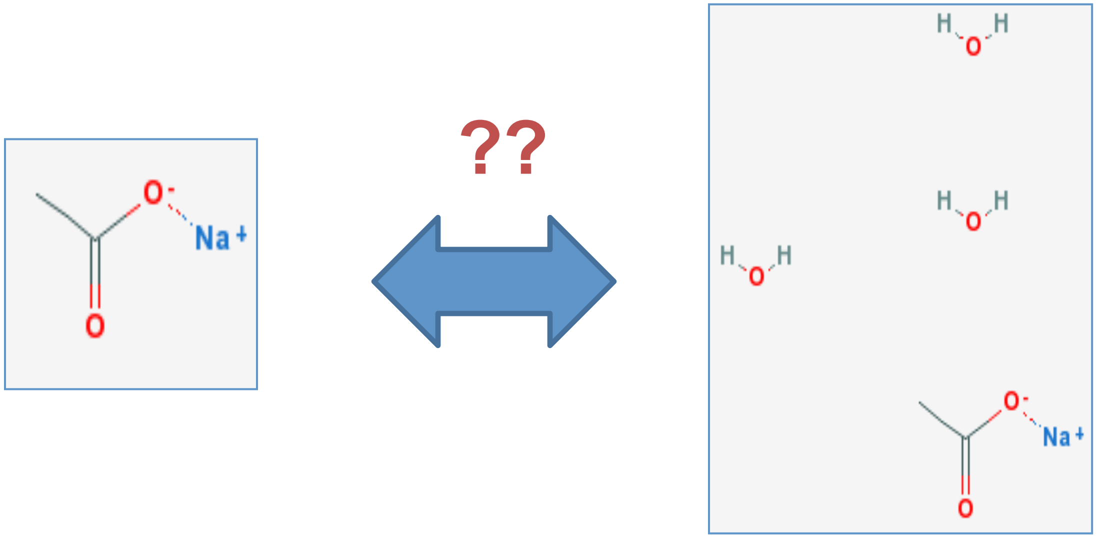


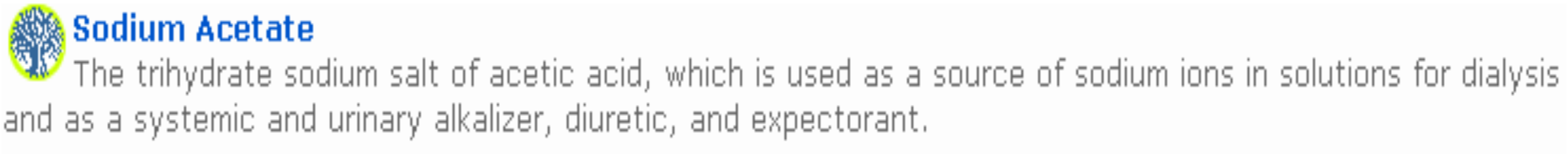




Salt-form drawing variations are common

What do you mean by “sodium acetate”?




Sodium Acetate

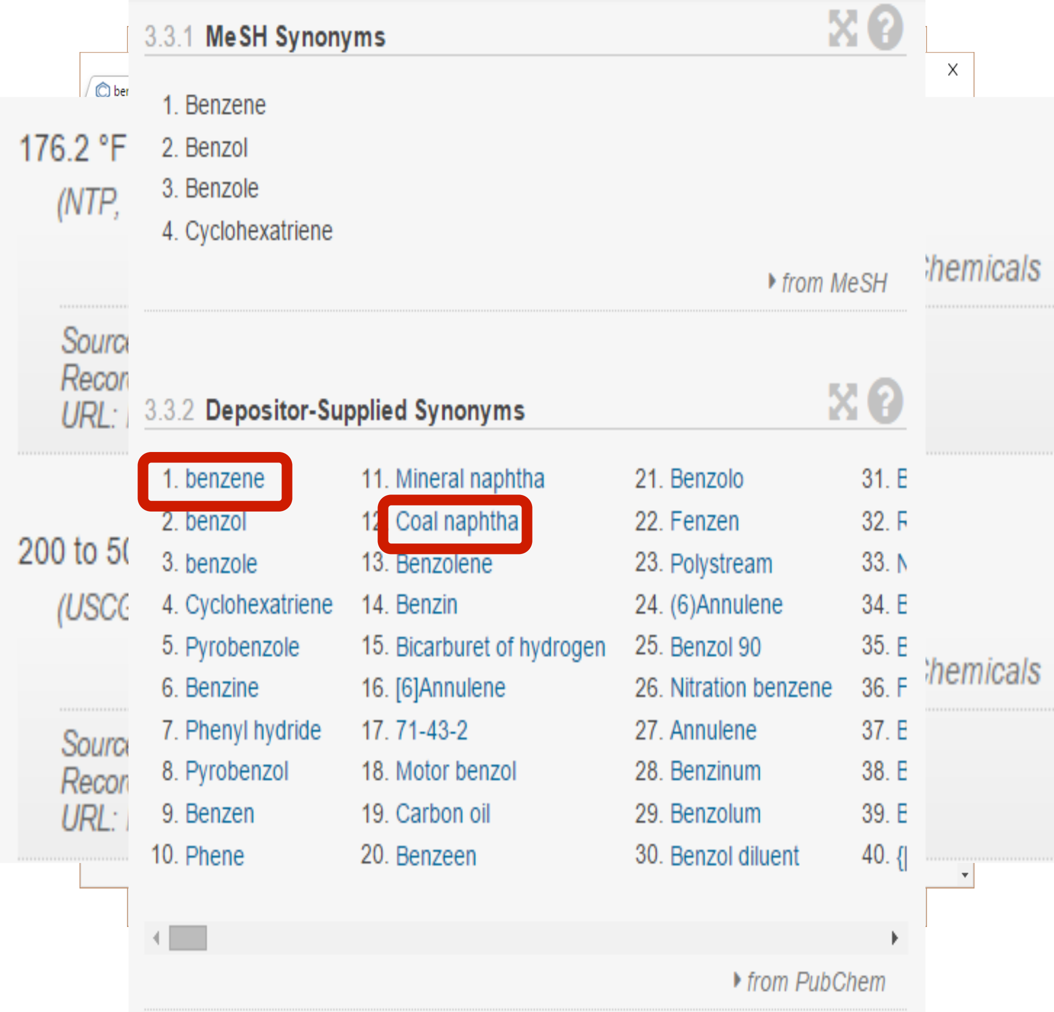
The trihydrate sodium salt of acetic acid, which is used as a source of sodium ions in solutions for dialysis and as a systemic and urinary alkalizer, diuretic, and expectorant.

Chemical meaning of a substance may change upon context

Benzene boiling point case study

Benzene

Coal tar



176.2 °F

(NTP,

Sourc
Recor
URL:

200 to 5(

(USCG

Sourc
Recor
URL:

hemicals

hemicals

3.3.1 **MeSH Synonyms**

1. Benzene
2. Benzol
3. Benzole
4. Cyclohexatriene

from MeSH

3.3.2 **Depositor-Supplied Synonyms**

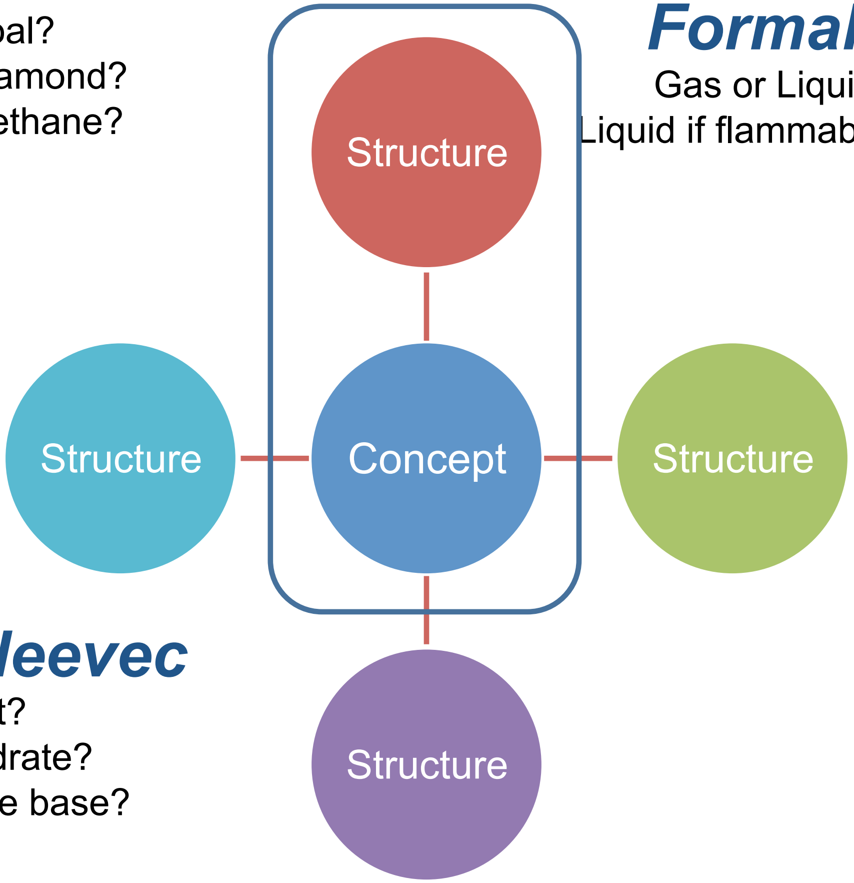
1. benzene
2. benzol
3. benzole
4. Cyclohexatriene
5. Pyrobenzole
6. Benzine
7. Phenyl hydride
8. Pyrobenzol
9. Benzen
10. Phene
11. Mineral naphtha
12. Coal naphtha
13. Benzolene
14. Benzin
15. Bicarburet of hydrogen
16. [6]Annulene
17. 71-43-2
18. Motor benzol
19. Carbon oil
20. Benzeen
21. Benzolo
22. Fenzen
23. Polystream
24. (6)Annulene
25. Benzol 90
26. Nitration benzene
27. Annulene
28. Benzinum
29. Benzolum
30. Benzol diluent
31. E
32. F
33. N
34. E
35. E
36. F
37. E
38. E
39. E
40. {[

from PubChem

Many to many relationships

Carbon

Element?
Coal?
Diamond?
Methane?

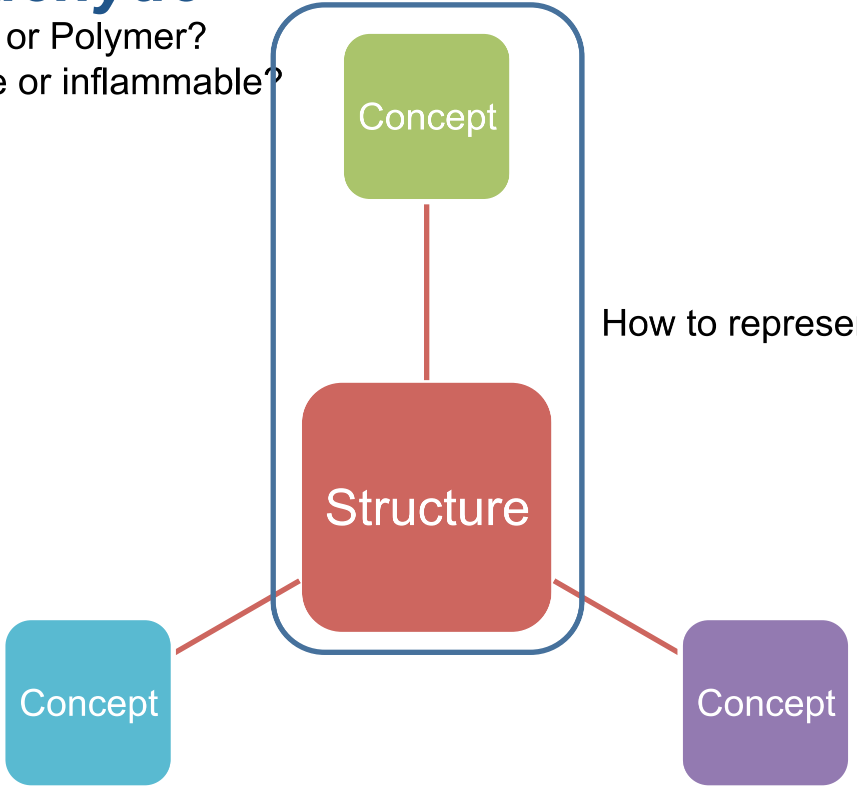


Formaldehyde

Gas or Liquid or Polymer?
Liquid if flammable or inflammable?

Gleevec

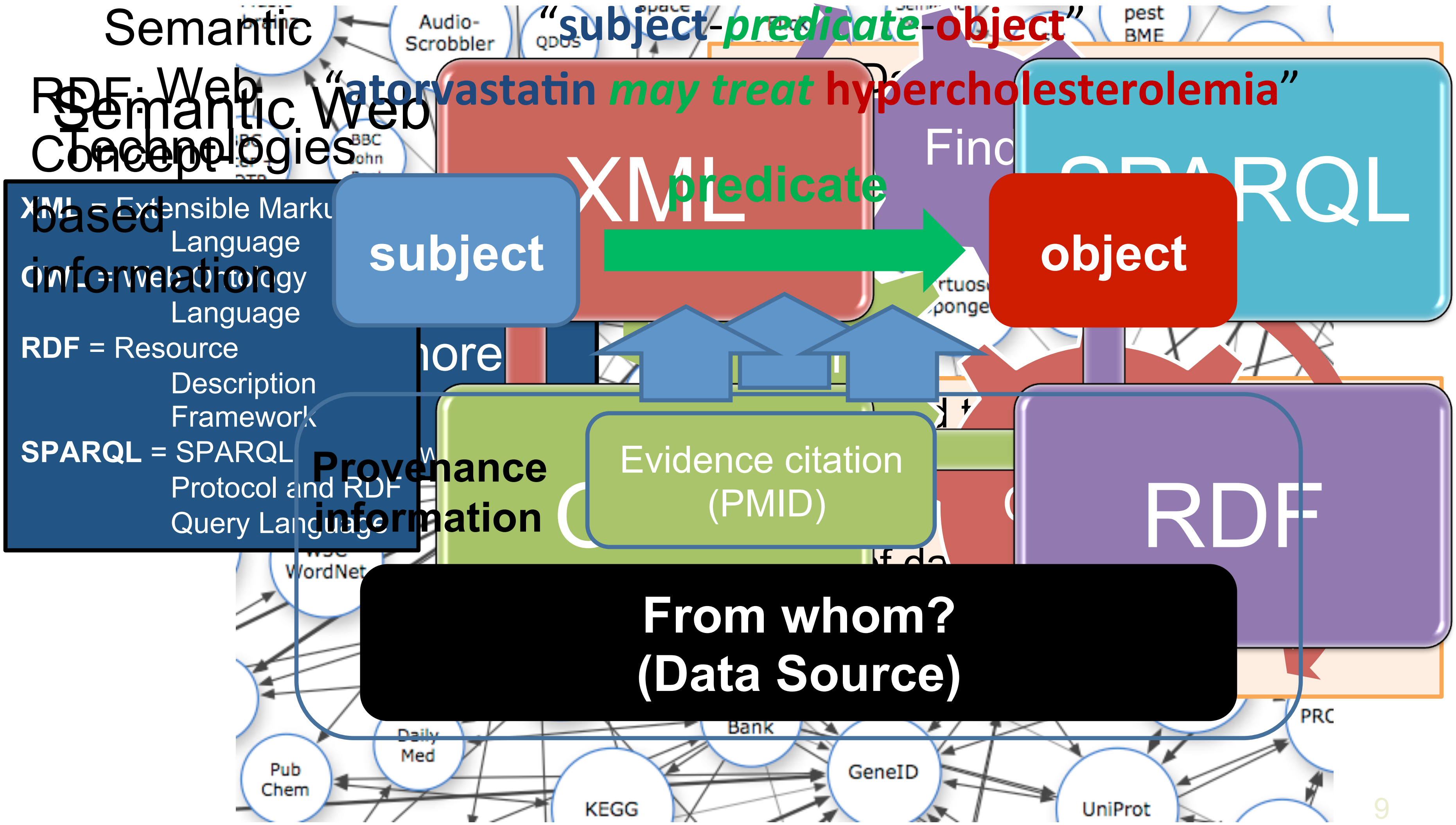
Salt?
Hydrate?
Free base?



How to represent?

Pick a preferred concept for a structure
Pick a preferred structure for a concept

Resource Description Framework (RDF) .. what is RDF?



Semantic Web Concept

Semantic Web Technologies

RDF: Web based information

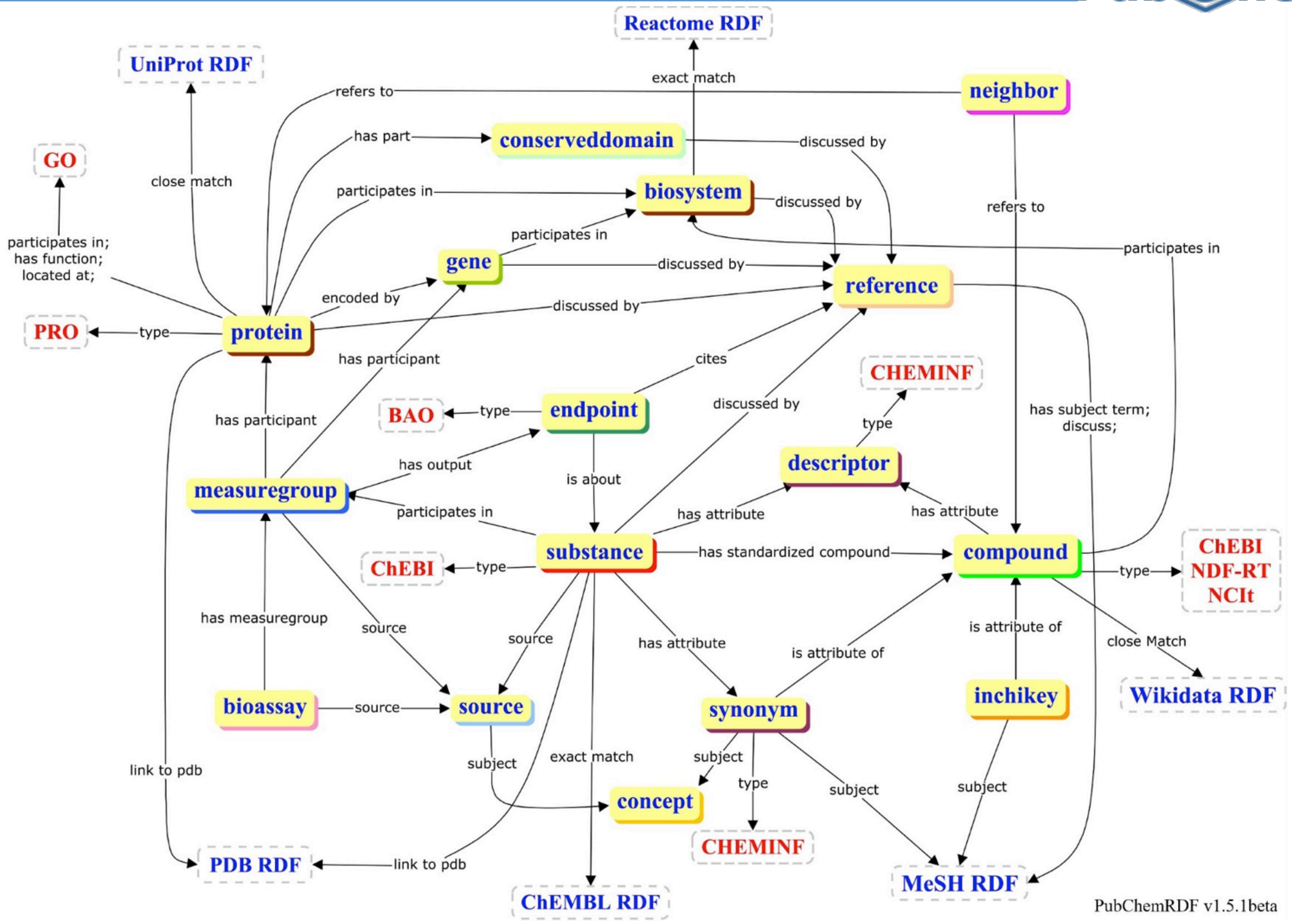
“subject-*predicate*-object”

“atorvastatin *may treat* hypercholesterolemia”

XML = Extensible Marku Language
OWL = Web Ontology Language
RDF = Resource Description Framework
SPARQL = SPARQL Protocol and RDF Query Language

Image credit: http://linkeddata.org/

PubChemRDF Overview



PubChemRDF v1.5.1beta

PubChemRDF Subdomains

Prefix	Namespace
compound	http://rdf.ncbi.nlm.nih.gov/pubchem/compound/
substance	http://rdf.ncbi.nlm.nih.gov/pubchem/substance/
descr	http://rdf.ncbi.nlm.nih.gov/pubchem/descriptor/
inchikey	http://rdf.ncbi.nlm.nih.gov/pubchem/inchikey/
syno	http://rdf.ncbi.nlm.nih.gov/pubchem/synonym/
bioassay	http://rdf.ncbi.nlm.nih.gov/pubchem/bioassay/
measuregroup	http://rdf.ncbi.nlm.nih.gov/pubchem/measuregroup/
endpoint	http://rdf.ncbi.nlm.nih.gov/pubchem/endpoint/
protein	http://rdf.ncbi.nlm.nih.gov/pubchem/protein/
conserveddomain	http://rdf.ncbi.nlm.nih.gov/pubchem/conserveddomain/
biosystem	http://rdf.ncbi.nlm.nih.gov/pubchem/biosystem/
gene	http://rdf.ncbi.nlm.nih.gov/pubchem/gene/
reference	http://rdf.ncbi.nlm.nih.gov/pubchem/reference/
nbr[a]	http://rdf.ncbi.nlm.nih.gov/pubchem/neighbor/
source	http://rdf.ncbi.nlm.nih.gov/pubchem/source/
concept	http://rdf.ncbi.nlm.nih.gov/pubchem/concept/
vocab	http://rdf.ncbi.nlm.nih.gov/pubchem/vocabulary#

http://pubchem.ncbi.nlm.nih.gov/rdf

PubChemRDF URIs

http://rdf.ncbi.nlm.nih.gov/pubchem/compound/CID727

http://rdf.ncbi.nlm.nih.gov/pubchem/substance/SID103554720

http://rdf.ncbi.nlm.nih.gov/pubchem/bioassay/AID1788

http://rdf.ncbi.nlm.nih.gov/pubchem/measuregroup/AID447528

http://rdf.ncbi.nlm.nih.gov/pubchem/protein/GI124375976

http://rdf.ncbi.nlm.nih.gov/pubchem/conserveddomain/PSSMID132758

http://rdf.ncbi.nlm.nih.gov/pubchem/gene/GID367

http://rdf.ncbi.nlm.nih.gov/pubchem/biosystem/BSID82991

http://rdf.ncbi.nlm.nih.gov/pubchem/reference/PMID10395478

http://rdf.ncbi.nlm.nih.gov/pubchem/inchikey/XUKUURHRXDUEBC-KAYWLYCHSA-N

http://rdf.ncbi.nlm.nih.gov/pubchem/synonym/MD5_9a05646d461669f86de312d88ab5748a

http://rdf.ncbi.nlm.nih.gov/pubchem/concept/ATC_L01XE

http://rdf.ncbi.nlm.nih.gov/pubchem/source/ChEMBL

PubChemRDF Composite URIs

http://rdf.ncbi.nlm.nih.gov/pubchem/descriptor/CID727_LogP_1

http://rdf.ncbi.nlm.nih.gov/pubchem/descriptor/CID727_LogP_2

http://rdf.ncbi.nlm.nih.gov/pubchem/measuregroup/AID1788_1

http://rdf.ncbi.nlm.nih.gov/pubchem/measuregroup/AID363_PMID16161995

http://rdf.ncbi.nlm.nih.gov/pubchem/endpoint/SID103164874_AID443491

http://rdf.ncbi.nlm.nih.gov/pubchem/endpoint/SID99445338_AID2202_1

http://rdf.ncbi.nlm.nih.gov/pubchem/endpoint/SID8033500_AID363_PMID10395478

http://rdf.ncbi.nlm.nih.gov/pubchem/protein/GI2506129GI254763435

http://rdf.ncbi.nlm.nih.gov/pubchem/neighbor/CID60823_CID68019409_2DSimilarity

http://rdf.ncbi.nlm.nih.gov/pubchem/neighbor/CID60823_CID68019409_2DTanimotoScore

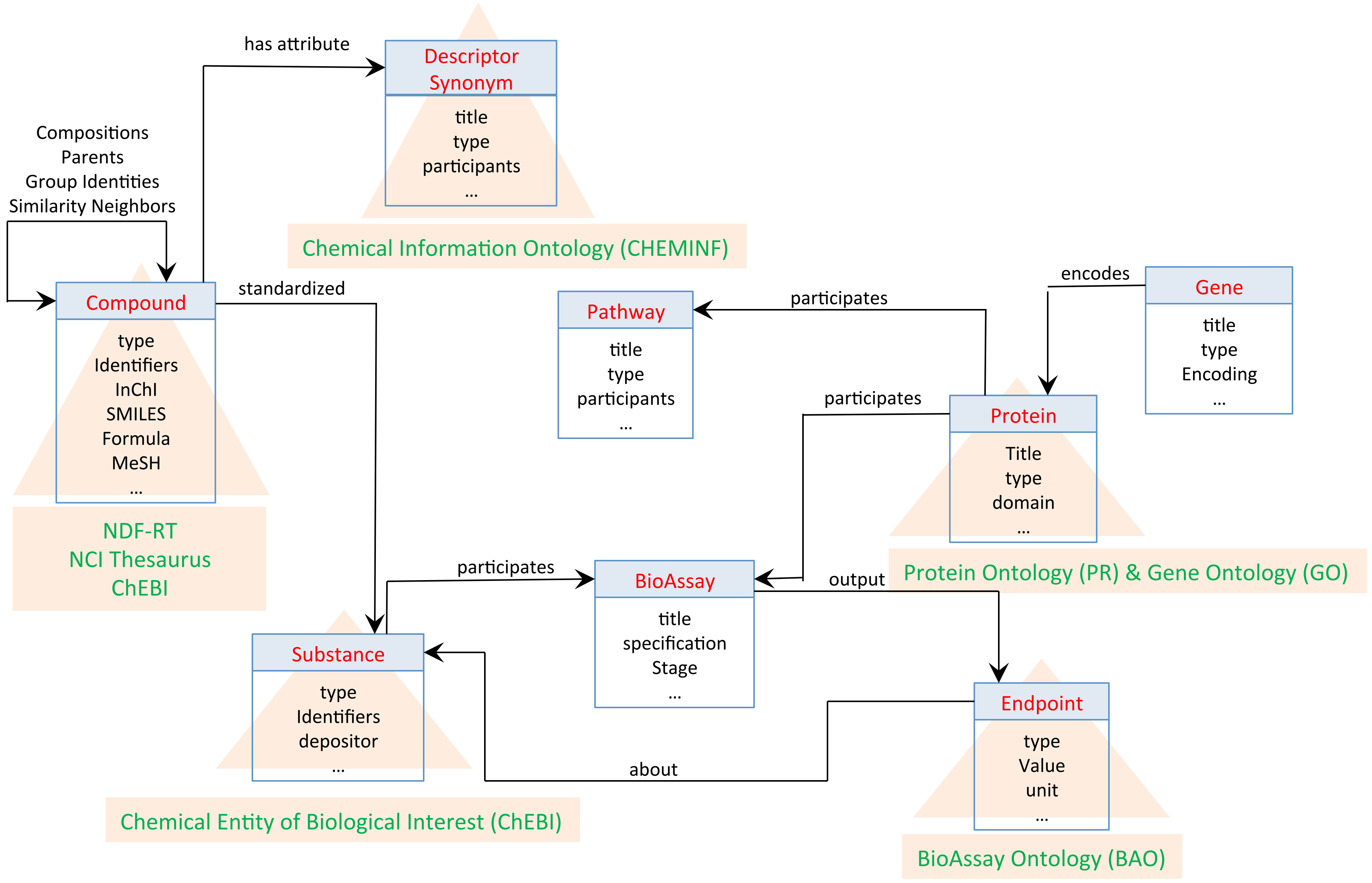
http://rdf.ncbi.nlm.nih.gov/pubchem/neighbor/CID60823_CID11330946_3DSimilarity

http://rdf.ncbi.nlm.nih.gov/pubchem/neighbor/CID60823_CID11330946_3DShapeTanimotoScore

http://rdf.ncbi.nlm.nih.gov/pubchem/neighbor/CID60823_CID11330946_3DFeatureTanimotoScore

PubChemRDF Overview: ontology-based data integration

PubChem

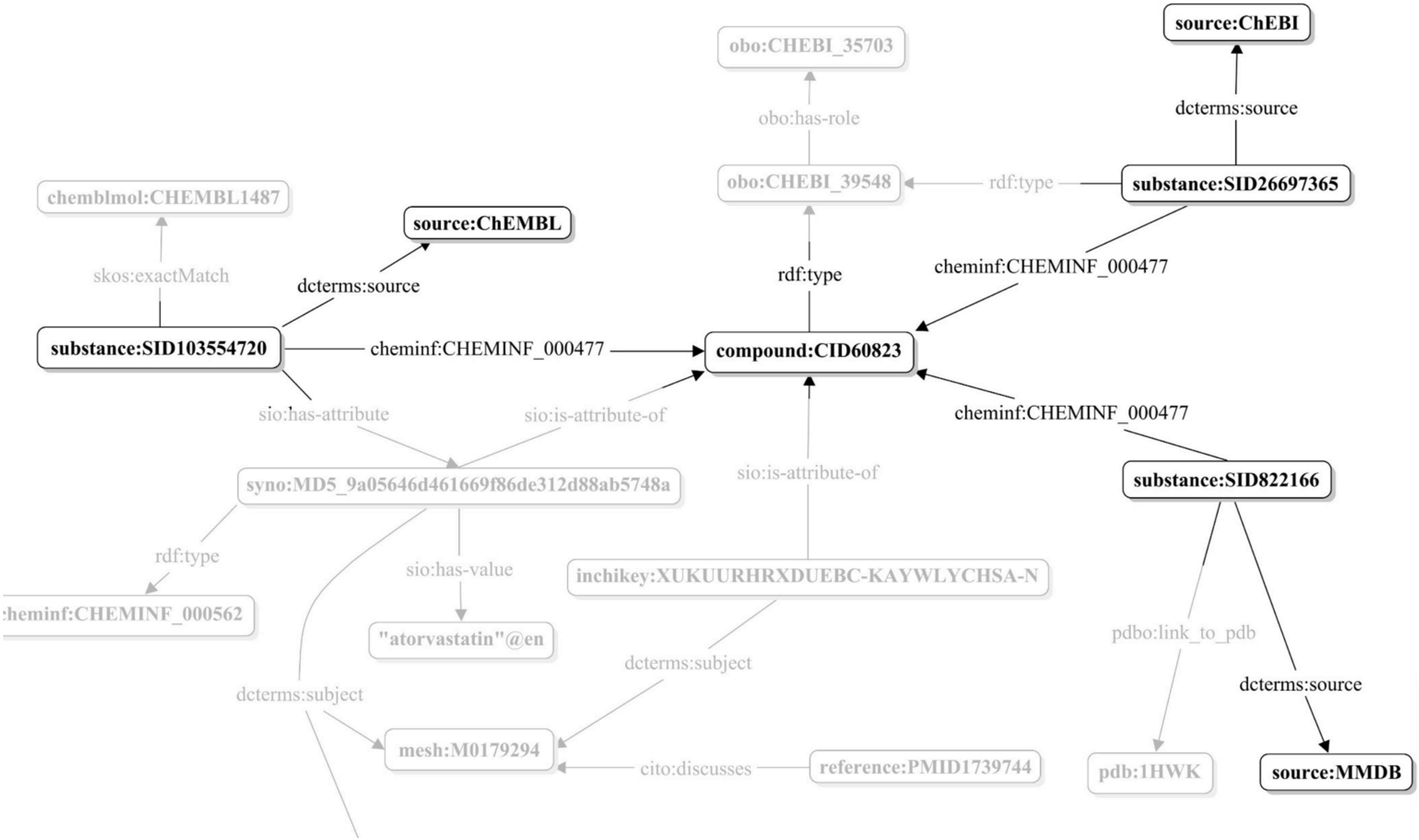


PubChemRDF Ontologies

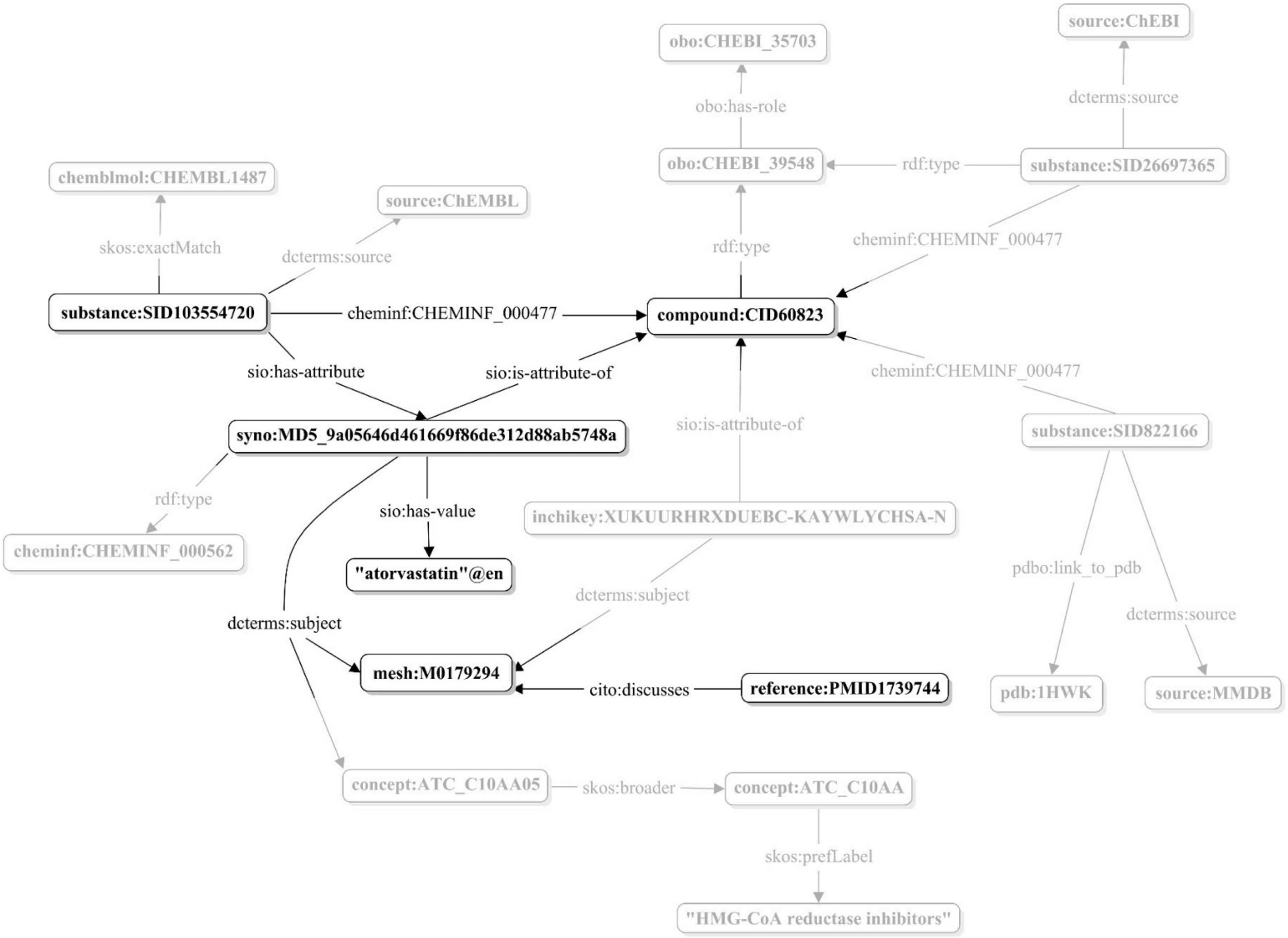
Prefix	Namespace	Vocabularies
rdfs	http://www.w3.org/2000/01/rdf-schema#	RDF Schema
rdf	http://www.w3.org/1999/02/22-rdf-syntax-ns#	RDF
owl	http://www.w3.org/2002/07/owl#	OWL
xsd	http://www.w3.org/2001/XMLSchema#	XML Schema
ndfrt	http://evs.nci.nih.gov/ftp1/NDF-RT/NDF-RT.owl#	NDF-RT
ncit	http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#	NCIt
sio[a]	http://semanticscience.org/resource/	SIO
cheminf[a]	http://semanticscience.org/resource/	CHEMINF
skos	http://www.w3.org/2004/02/skos/core#	SKOS
obo	http://purl.obolibrary.org/obo/	BFO, OBI, IAO, UO, ChEBI, PR, GO
bao	http://www.bioassayontology.org/bao#	BAO
bp	http://www.biopax.org/release/biopax-level3.owl#	BioPAX
cito	http://purl.org/spar/cito/	CiTO
fabio	http://purl.org/spar/fabio/	FaBio
pdbo	http://rdf.wwpdb.org/schema/pdbx-v40.owl#	PDBo
dcterms	http://purl.org/dc/terms/	DCMI Terms
pav	http://purl.org/pav/	PAV
foaf	http://xmlns.com/foaf/0.1/	FOAF Vocabulary

Hierarchical Classifications

PubChemRDF Graph 1: compound aggregates substances from different sources

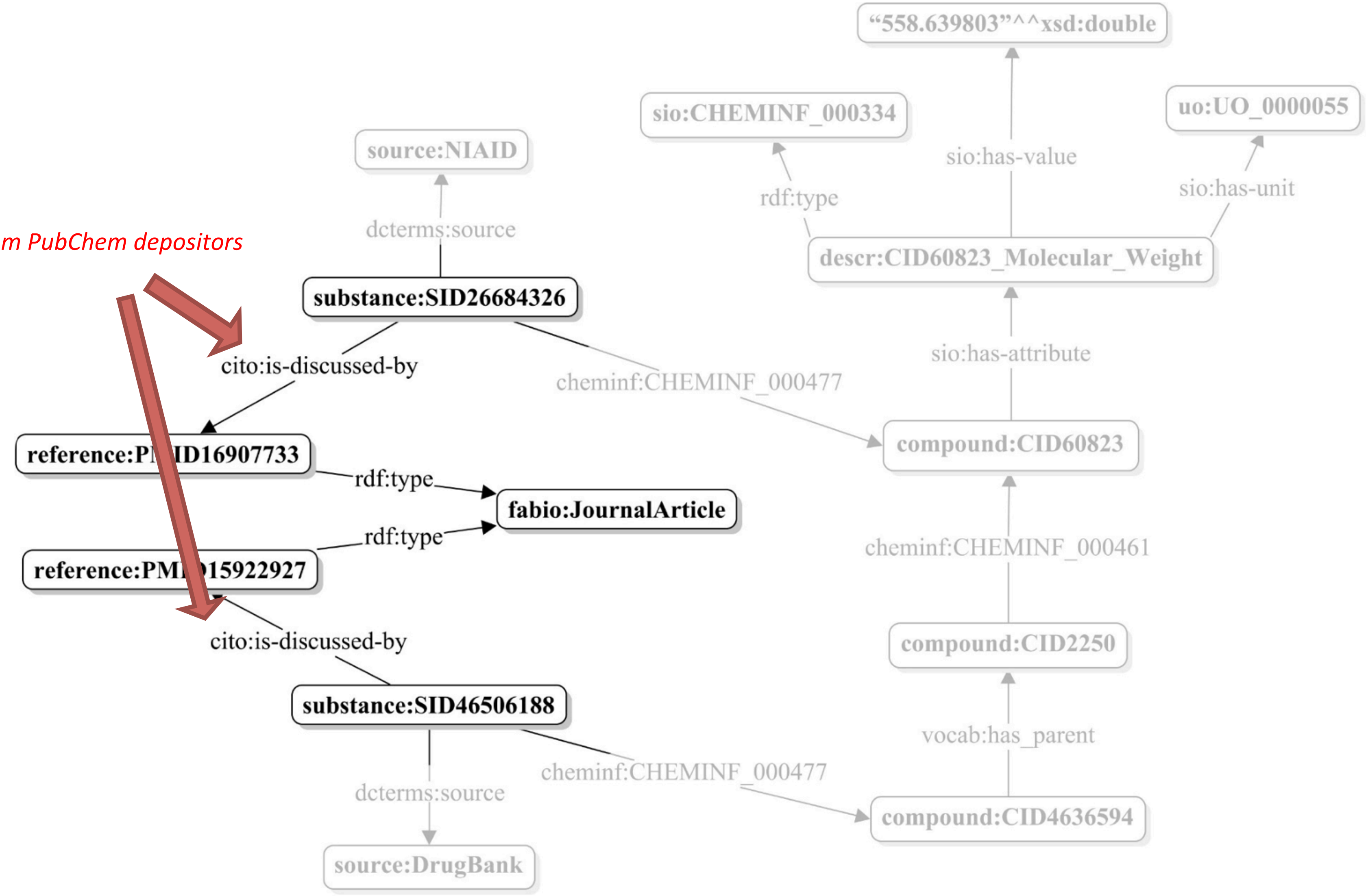


PubChemRDF Graph 2: synonym and MeSH annotations



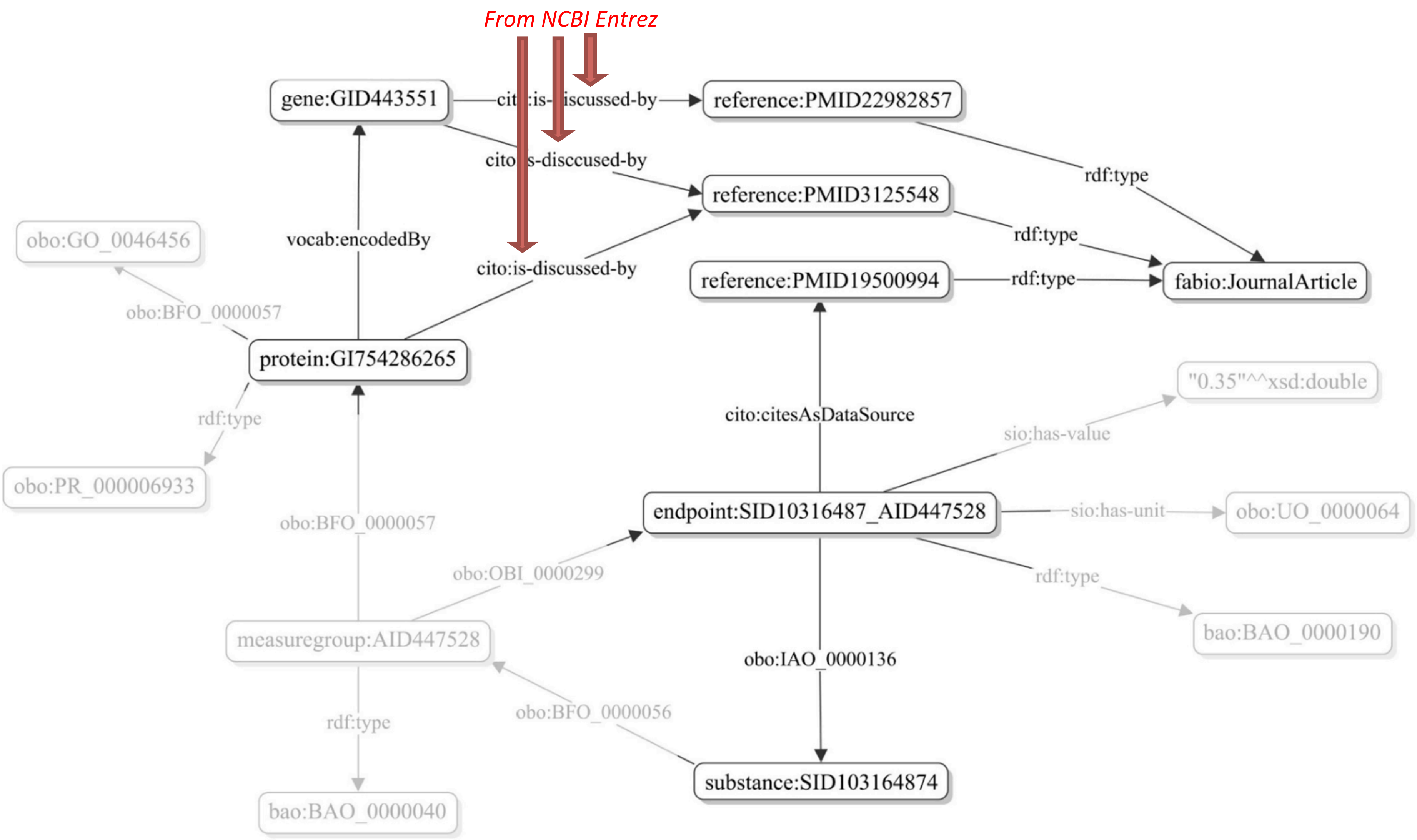
source:ChEBI
dcterms:source
obo:CHEBI_35703
obo:has-role
obo:CHEBI_39548
rdf:type
substance:SID26697365
chemblmol:CHEMBL1487
source:ChEMBL
skos:exactMatch
dcterms:source
rdf:type
cheminf:CHEMINF_000477
substance:SID103554720
cheminf:CHEMINF_000477
compound:CID60823
sio:has-attribute
sio:is-attribute-of
cheminf:CHEMINF_000477
syno:MD5_9a05646d461669f86de312d88ab5748a
sio:is-attribute-of
substance:SID822166
rdf:type
sio:has-value
cheminf:CHEMINF_000562
"atorvastatin"@en
inchikey:XUKUURHRXDUEBC-KAYWLYCHSA-N
pdbo:link_to_pdb
dcterms:subject
dcterms:source
dcterms:subject
mesh:M0179294
cito:discusses
reference:PMID1739744
pdb:1HWK
source:MMDB
concept:ATC_C10AA05
skos:broader
concept:ATC_C10AA
skos:prefLabel
"HMG-CoA reductase inhibitors"

PubChemRDF Graph 3: depositor-provided references for substances

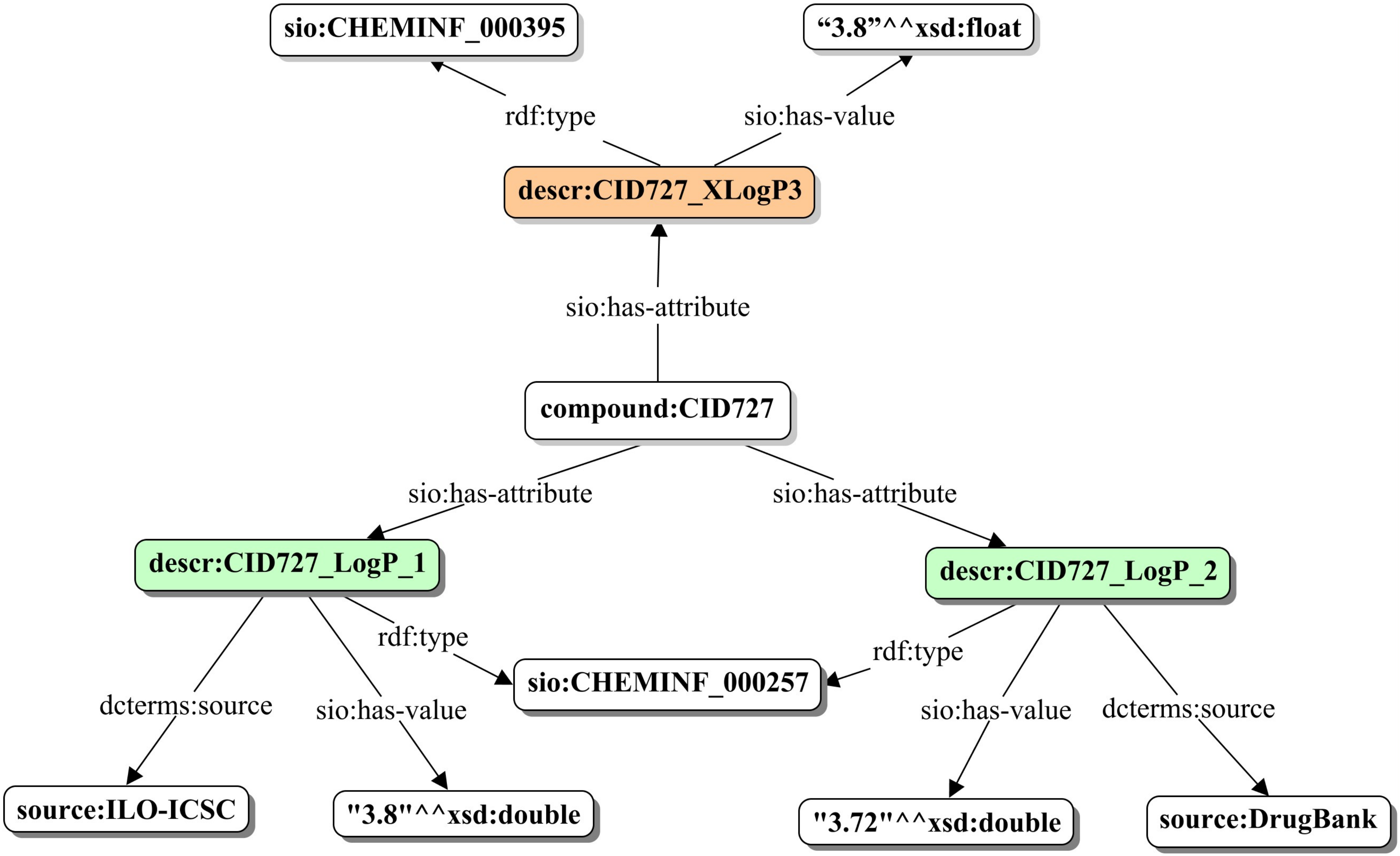


From PubChem depositors

PubChemRDF Graph 4: references for proteins, genes, bioassays, and so on



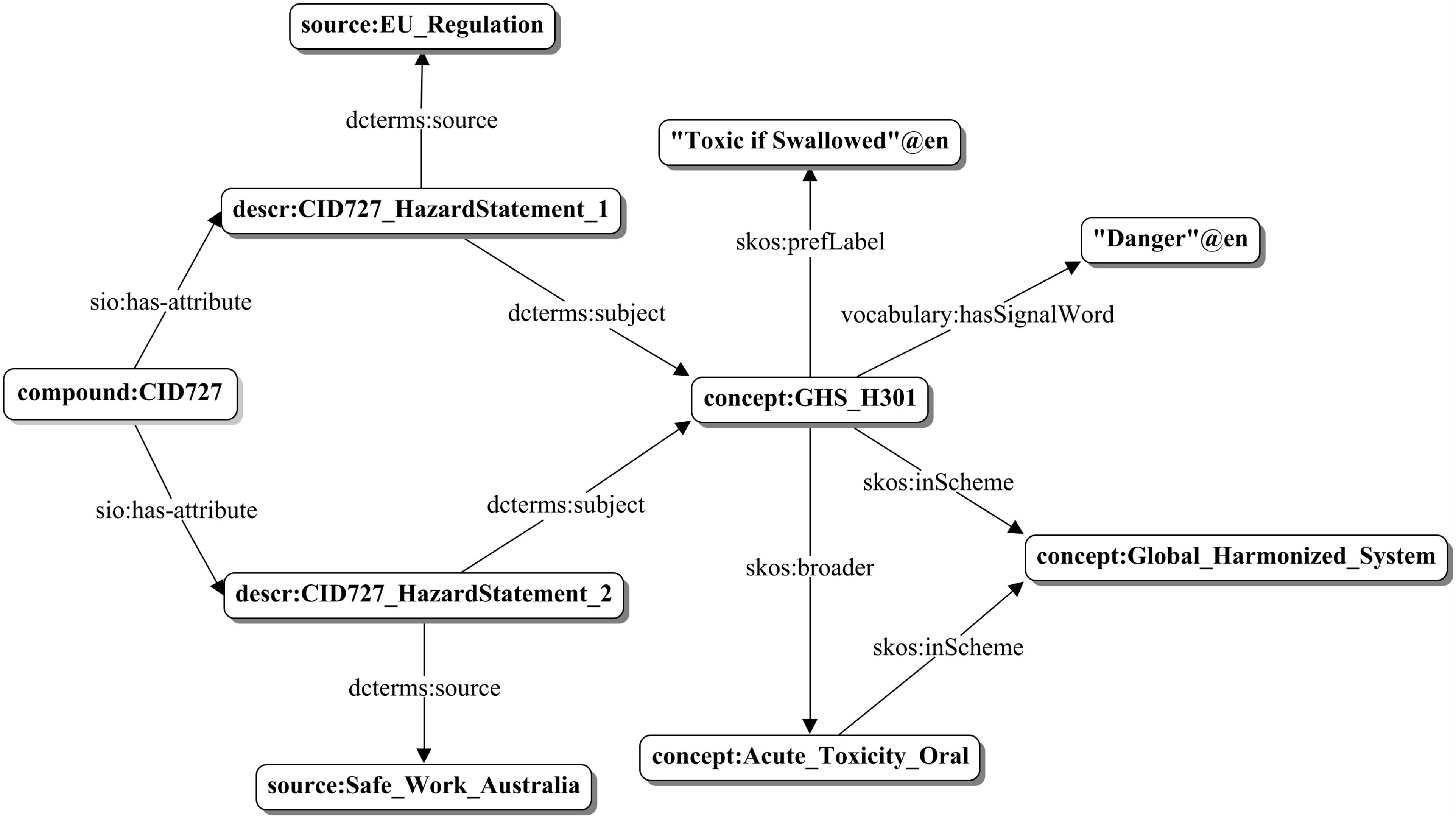
PubChemRDF Graph 5: Physical properties



PubChemRDF Ontologies for Physical Properties

Ontology	URIs	Properties
Chemical Information Ontology (CHEMINF)	sio:CHEMINF_000444	Auto-ignition Temperature
	sio:CHEMINF_000257	Boiling Point
	sio:CHEMINF_000443	Relative Evaporation Rate
	sio:CHEMINF_000417	Flash Point
	sio:CHEMINF_000191	Ionization Potential
	sio:CHEMINF_000436	Lower Explosive Limit
	sio:CHEMINF_000251	LogP
	sio:CHEMINF_000256	Melting Point
	sio:CHEMINF_000441	Odor Threshold
	sio:CHEMINF_000442	pH
	sio:CHEMINF_000435	Upper Explosive Limit
	sio:CHEMINF_000440	Vapor Density
	sio:CHEMINF_000255	Vapor Pressure
Chemical Methods Ontology (CHMO)	obo:CHMO_0001487	Decomposition
	obo:CHMO_0002818	Optical Rotation
	obo:CHMO_0002815	Solubility
Phenotypic Quality Ontology (PATO)	obo:PATO_0000014	Color
	obo:PATO_0001019	Density
	obo:PATO_0001884	Hydrophobicity
	obo:PATO_0000058	Odor
	obo:PATO_0001461	Surface Tension

PubChemRDF Graph 6: Globally Harmonized System Statements



PubChemRDF Synonym Classification

CAS RNs (*authoritative*):	513,833
CAS RNs (regex):	761,637
EC numbers (*authoritative*):	100,096
RTECS numbers (*authoritative*):	3,948
UN numbers (*authoritative*):	2,077
UN numbers (regex):	75
FDA UNIIs (*authoritative*):	47,313
FDA UNIIs (regex):	52,264
CHEBI IDs (*authoritative*):	45,411
CHEBI IDs (regex):	2,360
Drug Trade Names:	24,163
WHO INN names:	63,250
FDA UNII names:	154,250
EPA SRS synonyms:	112,757
MESH terms:	167,800
NSC numbers (regex):	586,579
CHEMBL IDs (regex):	1,472,212
ZINC numbers (regex):	8,613,523
IUPAC names (OpenEye Lexichem computed):	15,579,470

PubChemRDF REST API: Content Negotiation

MIME Type	HTTP Accept Header	URI Suffix Extension
Abbreviated RDF/XML	application/rdf+xml+abbrev	rdfxml-abbrev
RDF/XML	application/rdf+xml text/rdf	rdfxml rdf xml
HTML	application/xhtml+xml text/html	html htm
TURTLE[a]	application/n3 application/rdf+n3 application/turtle application/x-turtle text/n3 text/turtle text/rdf+n3 text/rdf+turtle	turtle ttl n3
JSON[b]	application/json text/json	json
JSON-LD[c]	application/x-json+ld application/x-json+rdf application/json+ld application/json+rdf application/ld+json application/rdf+json	Jsonld Json-ld ldjson ld-json
N-TRIPLES	text/plain	ntriples (default)

New Format

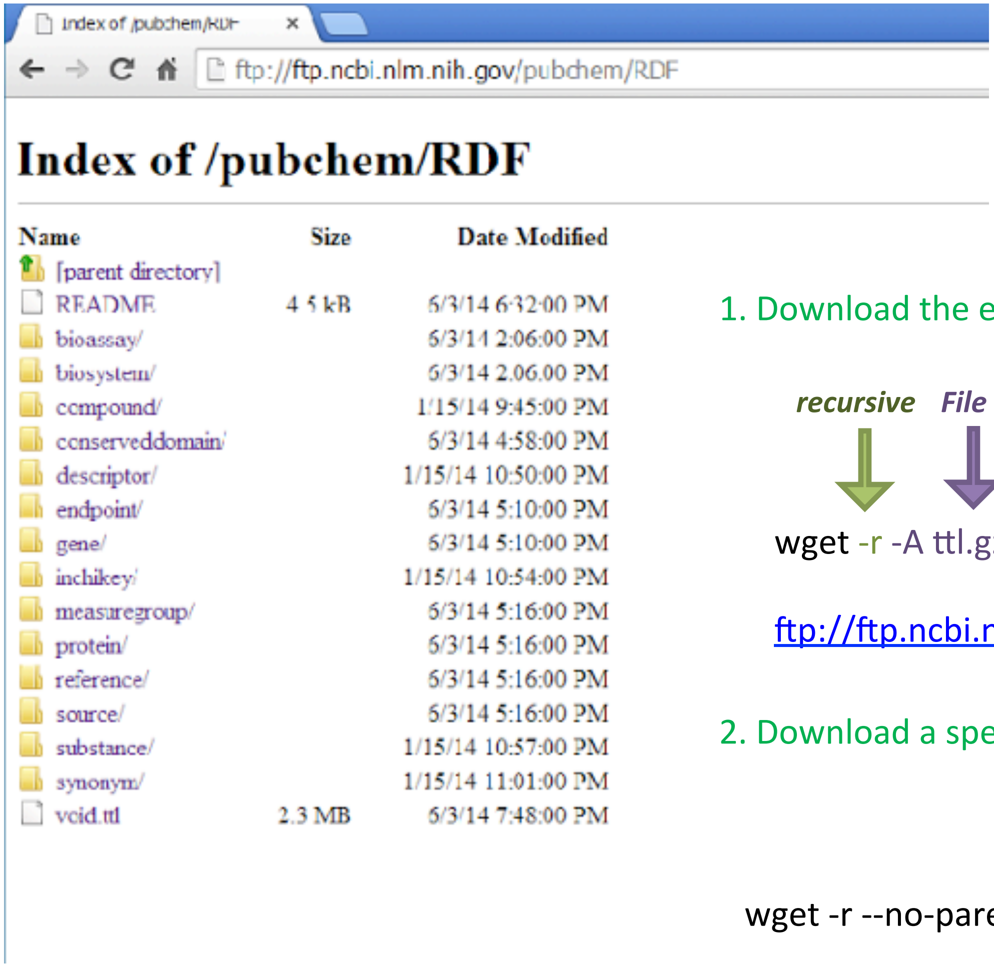
PubChemRDF REST API: Content Negotiation

- http://rdf.ncbi.nlm.nih.gov/pubchem/compound/CID2244.rdf
- http://rdf.ncbi.nlm.nih.gov/pubchem/compound/CID2244.xml
- http://rdf.ncbi.nlm.nih.gov/pubchem/compound/CID2244.rdfxml
- http://rdf.ncbi.nlm.nih.gov/pubchem/compound/CID2244.html
- http://rdf.ncbi.nlm.nih.gov/pubchem/compound/CID2244.turtle
- http://rdf.ncbi.nlm.nih.gov/pubchem/compound/CID2244.ttl
- http://rdf.ncbi.nlm.nih.gov/pubchem/compound/CID2244.json
- http://rdf.ncbi.nlm.nih.gov/pubchem/compound/CID2244.ntriples

Follow redirect
Content negotiation

```
curl -L -H "Accept: text/rdf"
http://rdf.ncbi.nlm.nih.gov/pubchem/compound/CID2244
```

PubChemRDF FTP Download



Index of /pubchem/RDF

Name	Size	Date Modified
[parent directory]		
README	4.5 kB	6/3/14 6:32:00 PM
bioassay/		6/3/14 2:06:00 PM
biosystem/		6/3/14 2:06:00 PM
compound/		1/15/14 9:45:00 PM
conserveddomain/		6/3/14 4:58:00 PM
descriptor/		1/15/14 10:50:00 PM
endpoint/		6/3/14 5:10:00 PM
gene/		6/3/14 5:10:00 PM
inchikey/		1/15/14 10:54:00 PM
measuregroup/		6/3/14 5:16:00 PM
protein/		6/3/14 5:16:00 PM
reference/		6/3/14 5:16:00 PM
source/		6/3/14 5:16:00 PM
substance/		1/15/14 10:57:00 PM
synonym/		1/15/14 11:01:00 PM
void.ttl	2.3 MB	6/3/14 7:48:00 PM

1. Download the entire directory of substance subdomain using ***wget***:

recursive *File suffix*

```
wget -r -A ttl.gz --no-host-directories
```

ftp://ftp.ncbi.nlm.nih.gov/pubchem/RDF/substance

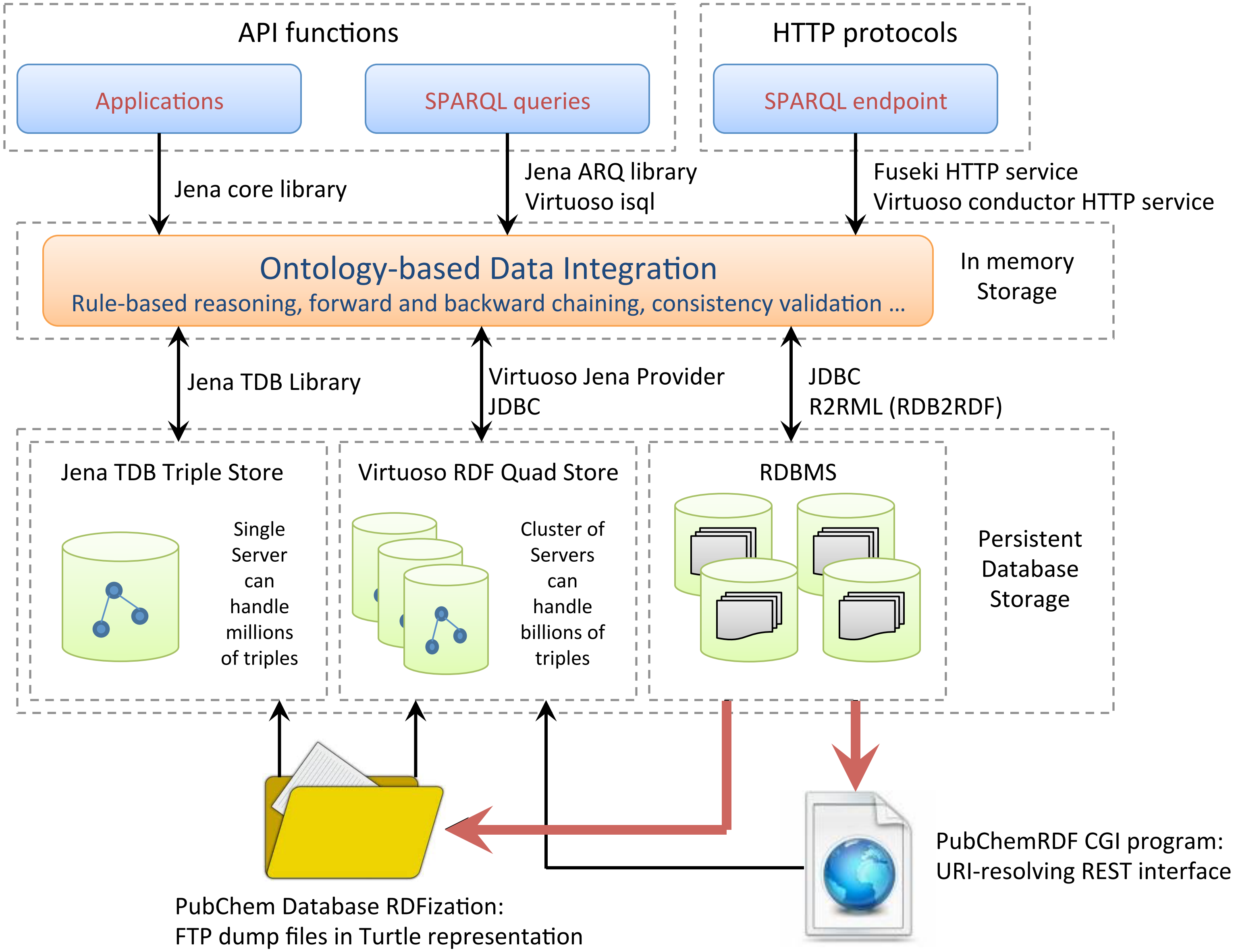
2. Download a specific type of link (substance to compound):

File suffix

```
wget -r --no-parent -A 'pc_substance2compound_*.ttl.gz'
```

ftp://ftp.ncbi.nlm.nih.gov/pubchem/RDF/substance

PubChemRDF Utility



PubChemRDF Use Cases: SPARQL query

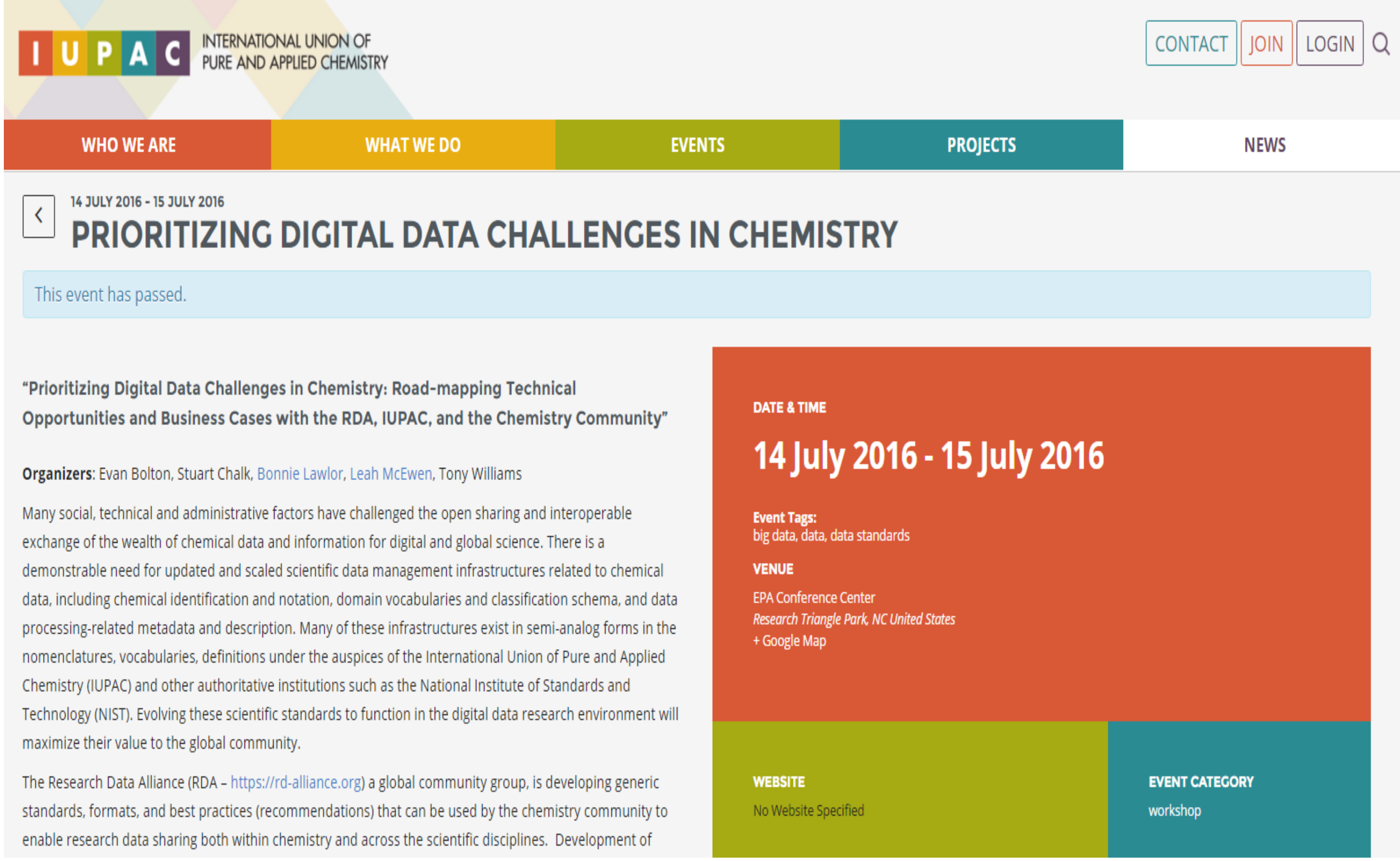
Q: What adverse effects of chemicals that are oral acute toxic according to GHS statement have been reported in PubMed literature, annotated by MeSH indexing?

```
PREFIX cito: <http://purl.org/spar/cito/>
PREFIX fabio: <http://purl.org/spar/fabio/>
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
PREFIX sio: <http://semanticscience.org/resource/>
PREFIX dcterms: <http://purl.org/dc/terms/>
PREFIX meshv: <http://id.nlm.nih.gov/mesh/vocab#>
PREFIX mesh: <http://id.nlm.nih.gov/mesh/>

select distinct ?disease ?diseaselabel
where {
  ?compound sio:has-attribute/dcterms:subject/skos:broader/concept:Acute_Toxicity_Oral .
  ?syno sio:is-attribute-of ?compound .
  ?syno dcterms:subject ?meshconcept .
  ?pmid cito:discusses ?meshconcept .
  ?pmid fabio:hasSubjectTerm ?DQpair .
  ?DQpair meshv:hasQualifier mesh:Q000009 .
  ?pmid cito:discusses ?disease .
  ?disease rdf:type meshv:SCR_Disease .
  ?disease rdfs:label ?diseaselabel .
}
```

Image credit:
http://msmissmrs.co.uk/wp-content/uploads/2015/06/community-words.jpg

RDA/IUPAC Workshop at EPA



14 JULY 2016 - 15 JULY 2016

PRIORITIZING DIGITAL DATA CHALLENGES IN CHEMISTRY

This event has passed.

"Prioritizing Digital Data Challenges in Chemistry: Road-mapping Technical Opportunities and Business Cases with the RDA, IUPAC, and the Chemistry Community"

Organizers: Evan Bolton, Stuart Chalk, Bonnie Lawlor, Leah McEwen, Tony Williams

Many social, technical and administrative factors have challenged the open sharing and interoperable exchange of the wealth of chemical data and information for digital and global science. There is a demonstrable need for updated and scaled scientific data management infrastructures related to chemical data, including chemical identification and notation, domain vocabularies and classification schema, and data processing-related metadata and description. Many of these infrastructures exist in semi-analog forms in the nomenclatures, vocabularies, definitions under the auspices of the International Union of Pure and Applied Chemistry (IUPAC) and other authoritative institutions such as the National Institute of Standards and Technology (NIST). Evolving these scientific standards to function in the digital data research environment will maximize their value to the global community.

The Research Data Alliance (RDA – https://rd-alliance.org) a global community group, is developing generic standards, formats, and best practices (recommendations) that can be used by the chemistry community to enable research data sharing both within chemistry and across the scientific disciplines. Development of

DATE & TIME

14 July 2016 - 15 July 2016

Event Tags:
big data, data, data standards

VENUE

EPA Conference Center
Research Triangle Park, NC United States
+ Google Map

WEBSITE

No Website Specified

EVENT CATEGORY

workshop

RDA/IUPAC Workshop at EPA

IUPAC Orange Book Ontology

(https://drive.google.com/open?id=1jRiJM048EyFfhE2u3ikI37wxIsG5rAaKZFirklNpA0g)

Develop a small scale ontology of chemical terms based on terms in IUPAC Orange Book as a case study. Foundational activities will look for example terminologies that have been converted to ontologies, identify where terms are currently being used and in what contexts, and look at relationships of those terms to others and potential differences in definitions. Terms will be transferred to a formal ontology in a plain bibliographic format, and a framework will be developed for augmenting the definition of terms to clarify the semantic meaning and context.

IUPAC Gold Book Data Structure

(https://drive.google.com/open?id=1hJdM7h90MBVLLUWBPtHe6cM-URJGi4zSlwYn8rXWNb8)

The IUPAC Gold Book is a valued compendium of terms sourcing from IUPAC published recommendations, including other Color Books and Pure and Applied Chemistry. The content is electronically accessible and linkable but not easily machine readable. This project is related to a current effort to extract the content data and term identifiers and migrate them into a more accessible and machine digestable format for increased usability.

Use Cases for Semantic Chemical Terminology Applications

(https://drive.google.com/open?id=1Ss5-qsIrgzSMTkcvEd52Iq-lwEYN2qCCN-BxN1ogGlI)

This scoping project will focus on researching the current chemical data transfer and communication landscape for potential applications of semantic terminology. Example use cases might include text books, patents, article and data indexing, standard protocols, experimental literature, published ontologies and thesauri with chemical terms, dictionaries for text mining, etc. Initial activities will analyze citations to terminology in the IUPAC Color Books (including the Gold Book) and Pure and Applied Chemistry.

You can help improve the state of the art

Image credit: http://www.idreamcareer.com/img/blog/1449649713-make-a-difference.jpg

Image credit: https://media.licdn.com/mpr/mpr/shrinknp_400_400/AAEAAQAAAAAAAAcPAAAAJDFjOTk4YzRiLWJiZTItNDBkNi1hYTYyLTJkOTFiZTBlNTMzMQ.jpg

Feel free to email me with questions and thoughts .. **evan_bolton@nih.gov**

Feel free to email me with questions and thoughts .. **evan_bolton@nih.gov**

We have infrastructure
We have data

We need volunteers to help!

Review terminology, provide use cases, perform assessments, help validate, and beyond.

Summary

PubChem

- PubChem RDF is intended for ontology-based data integration
- PubChem databases have been semantically exposed to linked open data
- REST interface can be accessed to resolve URI references
- FTP dump files can be bulk-loaded into open source triples stores
- LCSS information including physical properties and GSH statements have been added
- We need your help to make improvements

Feel free to email me with questions and thoughts .. **evan_bolton@nih.gov**

Acknowledgement

PubChem group

special thanks to ..

Yu Bo
Jie Chen
Siqian He
Jane He
Sunghwan Kim
Paul Thiessen
Leonid Zaslavsky

PubChem Collaborators

special thanks to ..

Colin Batchelor
Michel Dumontier
Janna Hastings
Hande Küçük
Stephan Schurer
Christopher Maloney
Roger Sayle
Daniel Lowe
Tony Williams
Chris Jakober

Thank you and Questions!