Maastricht University Leading in Learning!

Linked Open Data

Egon Willighagen (@egonwillighagen)

May 2012, Maastricht, HARMONY 2012

Open Data

- **Use**
- **Mix** (or convert!)
- **Reshare**

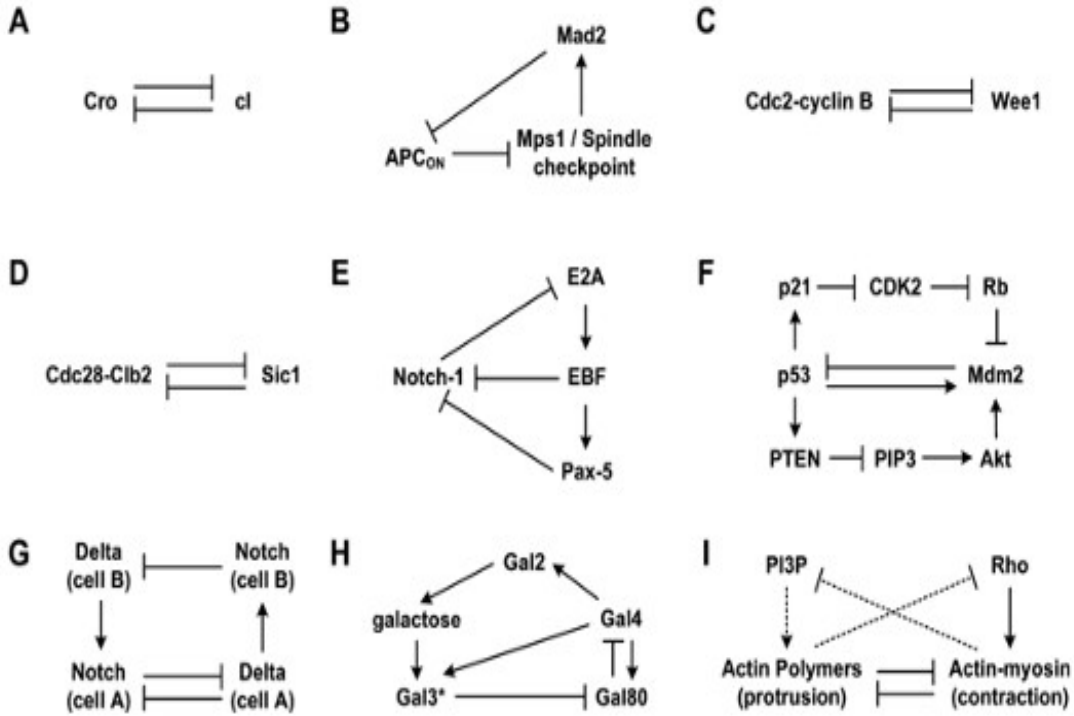
- Licenses
  - Creative Commons flavors, …
- Waivers
  - CCZero

Linked Data: meaning

- Why Semantic Web?
  - **Unique identifiers:** URLs (registries are redundant)
  - **Meaning:** what is glucose? What are Mad2, CDK2, Pax-5 (BioModel of the month 12/05, CCZero)?

Ontologies, …

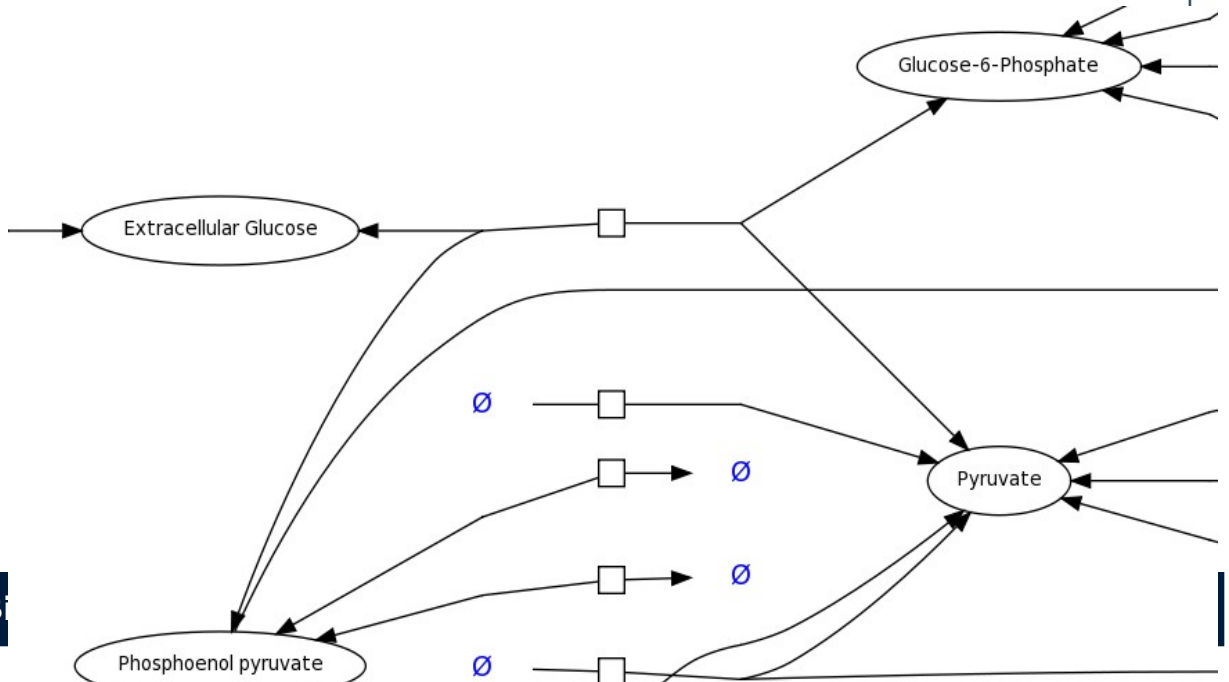
- Gene Ontology
- BioPortal, OLS, …
- …



Linked Data: linking

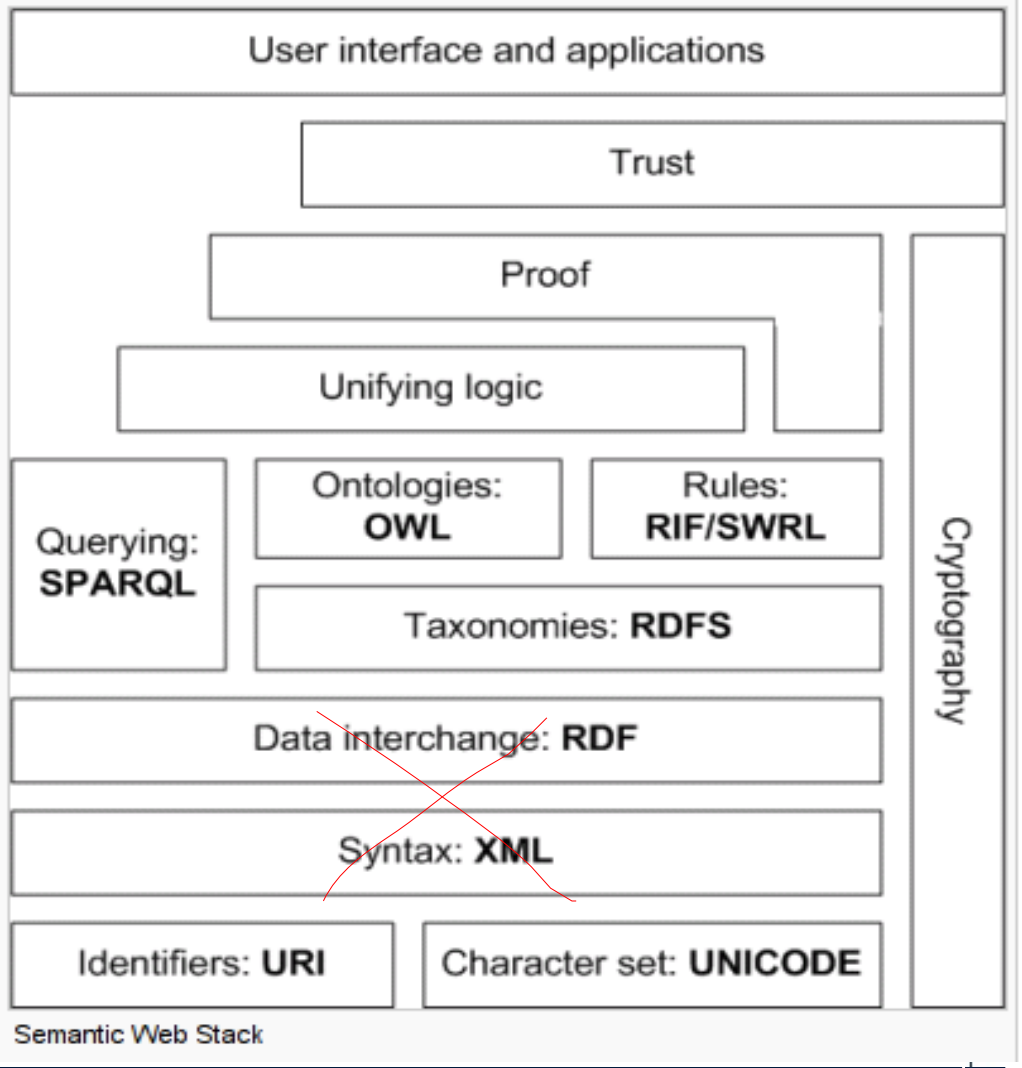
- Why Semantic Web?
  - **Web:** stand-up scientist
  - **Linking:** "Give me all models that predict something that predict glucose concentration"
  - "What are the chemical properties of glucose"

BIOMD0000000051
CCZero



Resource Description Framework

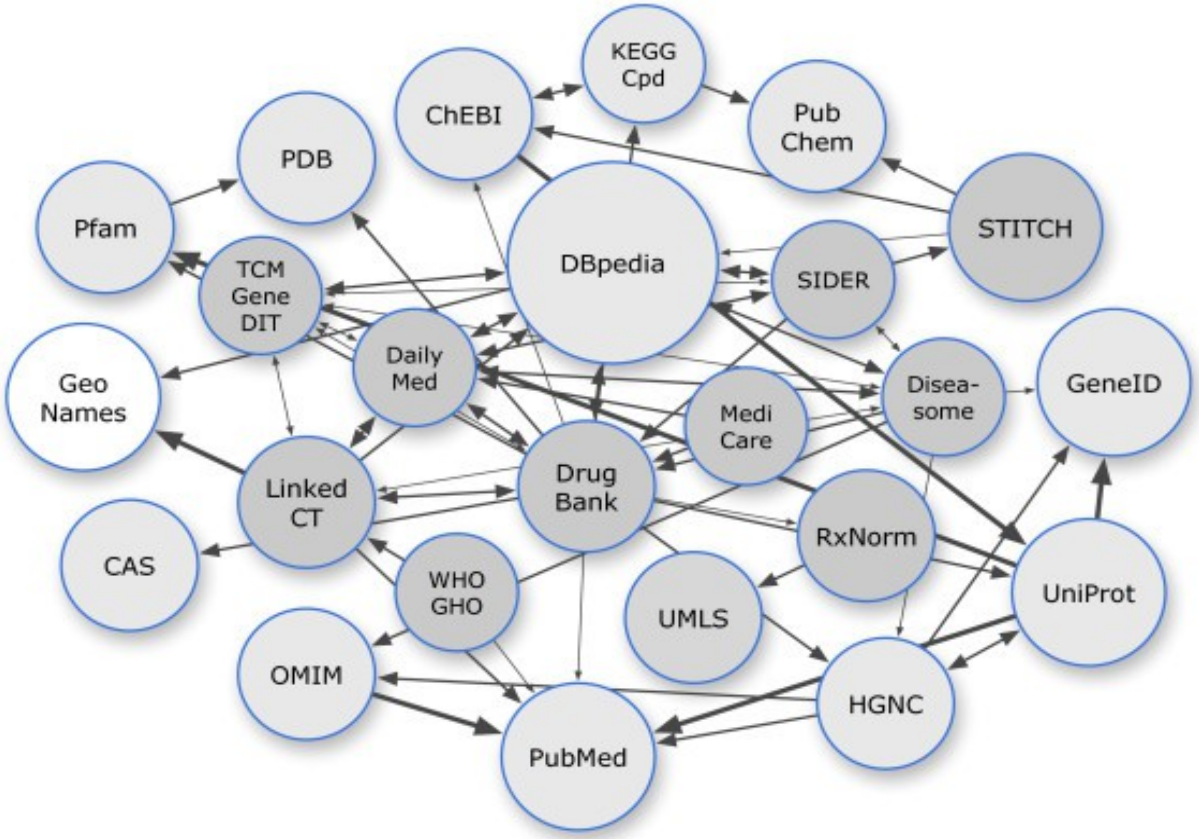
RDF != RDF/XML



Semantic Web Stack

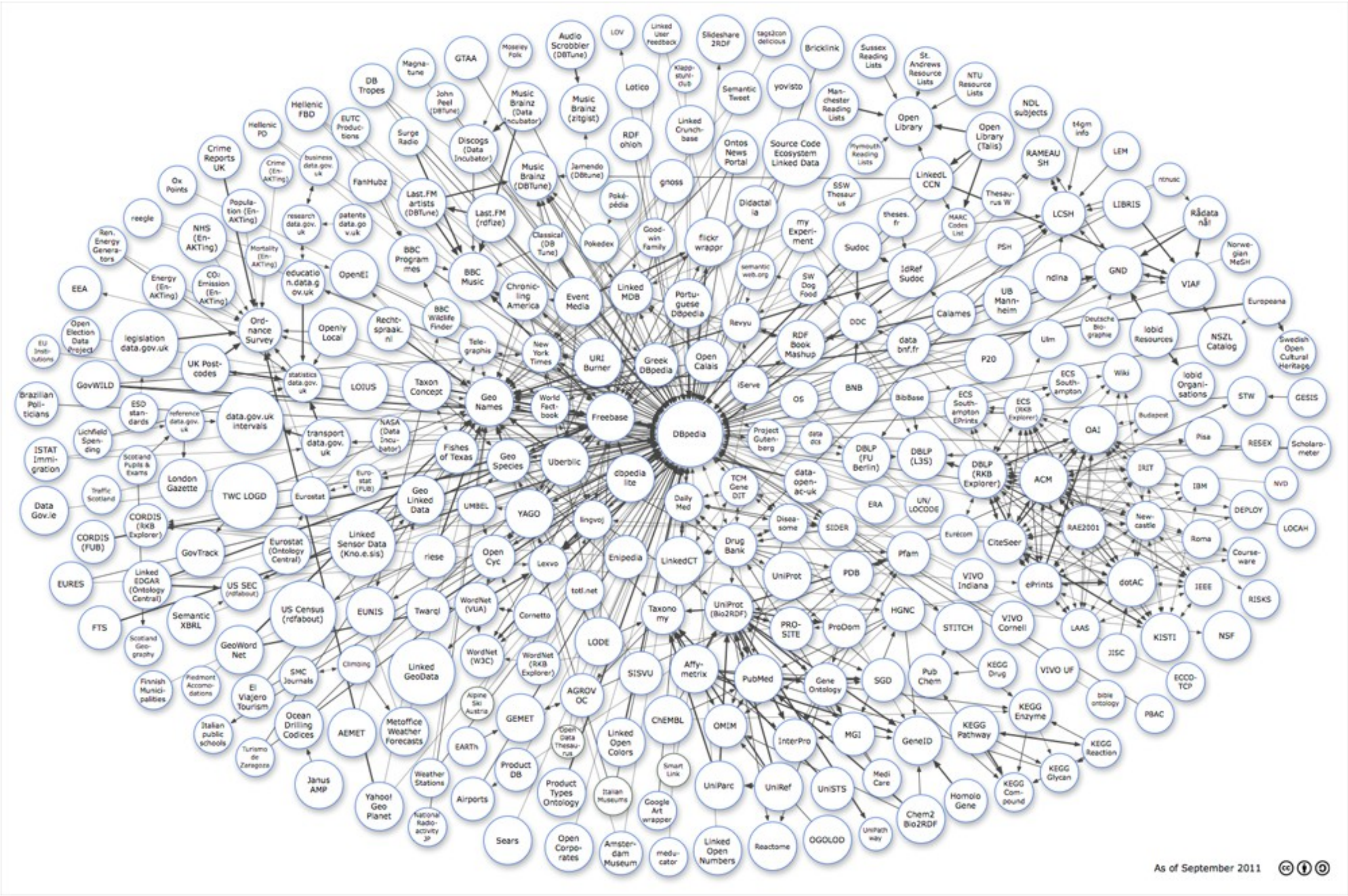
Wikipedia

Linked Open Drug Data



M. SamWald, et al, Linked open drug data for pharmaceutical research and development, 2011, JChemInf.

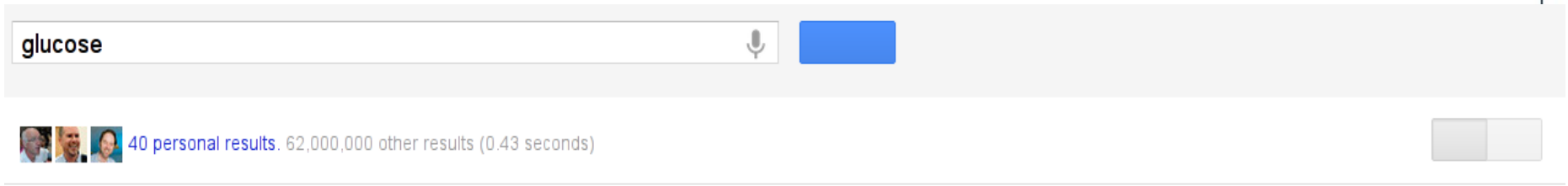
Linked Open Data



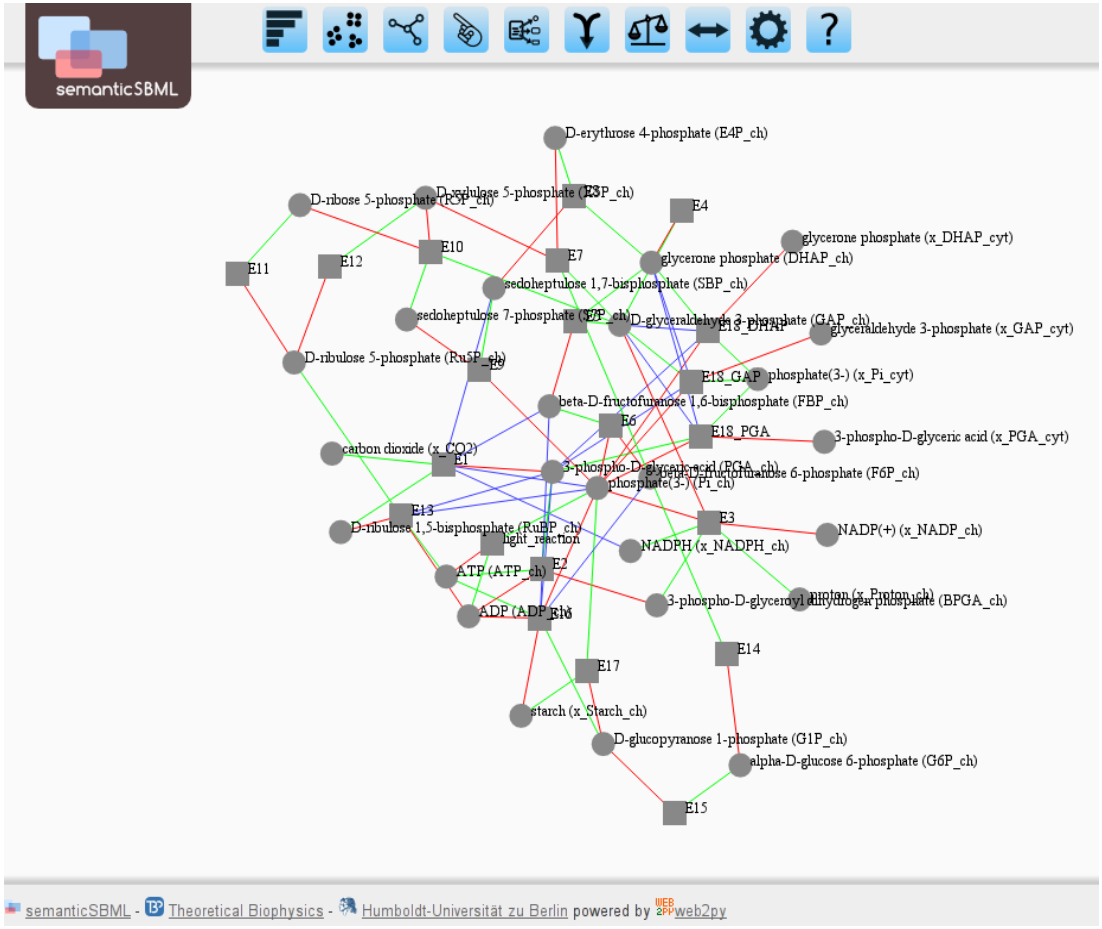
Linking Open Data cloud diagram, by Richard Cyganiak and Anja Jentzsch.
http://lod-cloud.net/ Sept 2011, CC-BY-SA.

Linked Open Data in the Life Sciences

Google ...


glucose
40 personal results. 62,000,000 other results (0.43 seconds)
Glucose - Wikipedia, the free encyclopedia
en.wikipedia.org/wiki/Glucose
Glucose (/'glu:koʊs/ or /-koʊz/; C6H12O6, also known as D-glucose, dextrose, or grape sugar) is a simple sugar (monosaccharide) and an important ...
Function - Structure and nomenclature - Physical properties - Production
Blood Glucose Levels: Testing and Normal Range
diabetes.webmd.com/blood-glucose
A blood glucose test measures the amount of a type of sugar, called glucose, in your blood. Glucose comes from carbohydrate foods.
What Is Glucose?
www.wisegeek.com/what-is-glucose.htm
6 days ago – Glucose is a type of sugar the body uses for energy. Diabetics, who have poorly regulated glucose levels, must be sure to...
Carbohydrates - Glucose
www.elmhurst.edu/~chm/vchembook/543glucose.html
Glucose is by far the most common carbohydrate and classified as a monosaccharide, an aldose, a hexose, and is a reducing sugar. It is also known as dextrose ...
Glucose Tests: The Test
labtestsonline.org/understanding/analytes/glucose/tab/test
23 Mar 2012 – Describes how glucose tests are used. when a glucose test is ordered.
Glucose
Glucose is a simple sugar and an important carbohydrate in biology. Cells use it as the primary source of energy and a metabolic intermediate. Glucose is one of the main products of photosynthesis and fuels for cellular respiration. Wikipedia
Formula: C6H12O6
Molar mass: 180.16 g/mol
Melting point: 146° C (294.8° F)
IUPAC ID: D-glucose
Density: 1.54 g/cm³
Soluble in: Water
Report a problem

W3C Health Care and Life Sciences interest group



Semantic Systems Biology

Protein-protein, protein-drug and gene-gene interacti
manipulation, it only takes into consideration one lay

Semantic web technologies enables the representatio

The systems biology task force will therefore identify
enabling the discovery of adverse effects when a pat

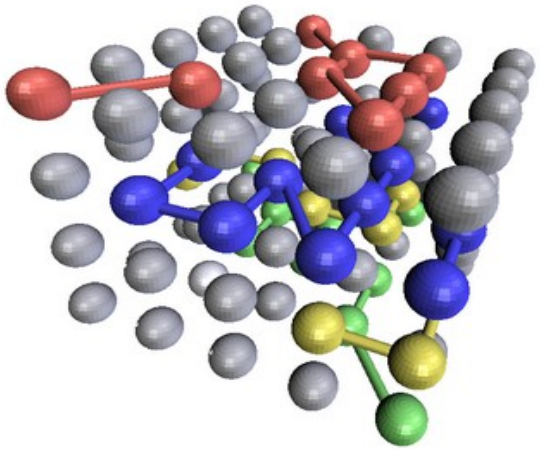


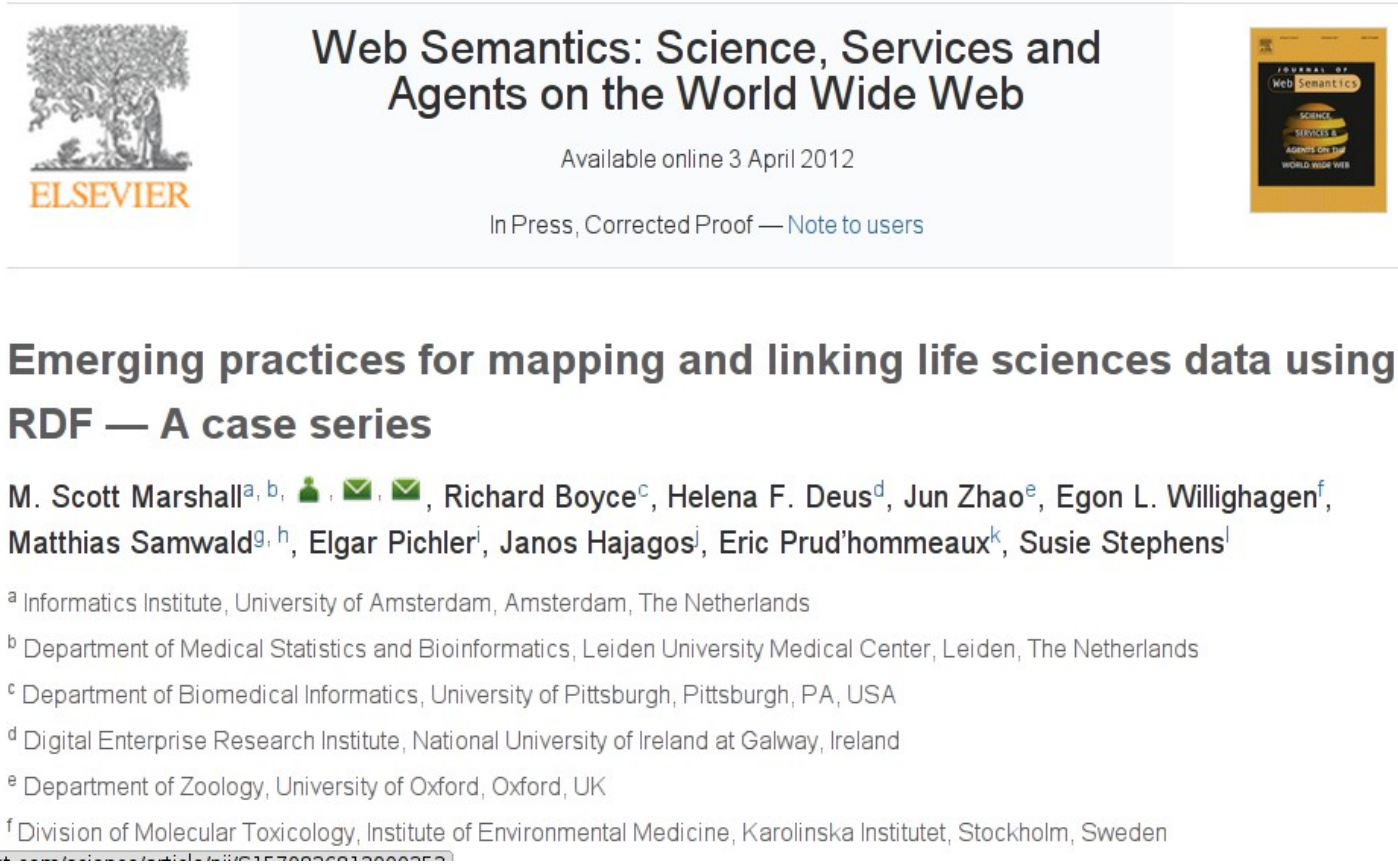
Figure 1. Different layers of information encode for di
based on cellular component (e.g. both gene product
others still may include information about co-disease

Initial Goals and Structure of the Task For

1. In this task force, we will pursue a pathway-based
MapK signalling pathway). This starts by 1) identifyin
pathway

2. The approach will be that of collecting the experim

W3C Health Care and Life Sciences interest group



ELSEVIER

Web Semantics: Science, Services and Agents on the World Wide Web

Available online 3 April 2012

In Press, Corrected Proof — Note to users

Emerging practices for mapping and linking life sciences data using RDF — A case series

M. Scott Marshall[a, b], Richard Boyce[c], Helena F. Deus[d], Jun Zhao[e], Egon L. Willighagen[f], Matthias Samwald[g, h], Elgar Pichler[i], Janos Hajagos[j], Eric Prud'hommeaux[k], Susie Stephens[l]

[a] Informatics Institute, University of Amsterdam, Amsterdam, The Netherlands

[b] Department of Medical Statistics and Bioinformatics, Leiden University Medical Center, Leiden, The Netherlands

[c] Department of Biomedical Informatics, University of Pittsburgh, Pittsburgh, PA, USA

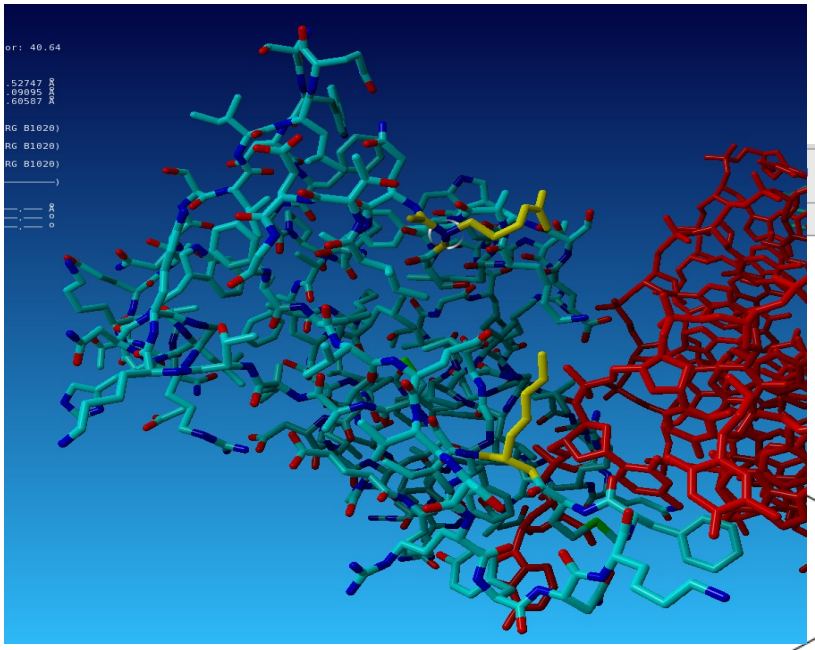
[d] Digital Enterprise Research Institute, National University of Ireland at Galway, Ireland

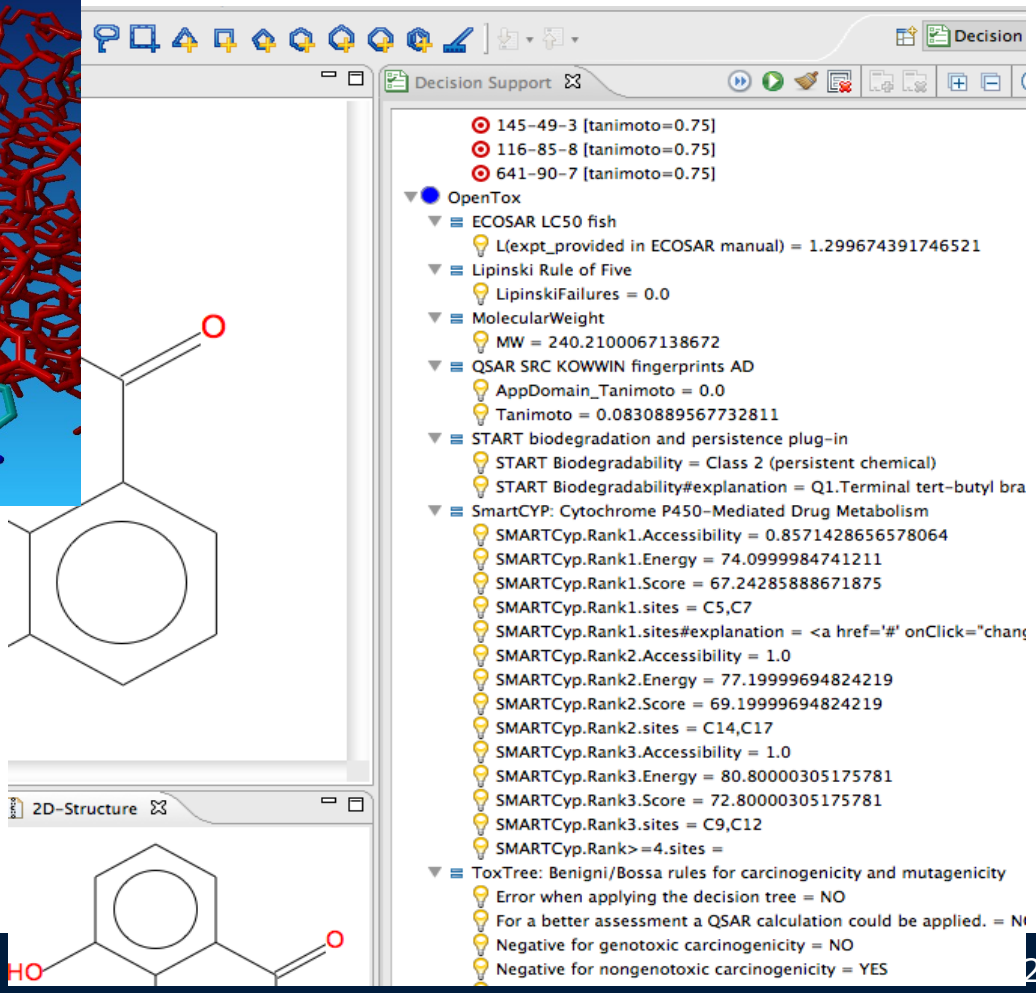
[e] Department of Zoology, University of Oxford, Oxford, UK

[f] Division of Molecular Toxicology, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden

http://thecostofknowledge.com/

Why does it matter again?





Conclusion

- Open Data
  - Use, mix, reshare
- Linked Data
  - Open standard (RDF, …)
    - … thus domain independent
  - Meaningful
    - via ontologies and resolvable, unique identifiers