Universität Rostock Traditio et Innovatio

SEMS simulation experiment management system

Introduction to ontologies in the life sciences

Dagmar Waltemath

dagmar.waltemath@uni-rostock.de

In the last fifty years the development of **new technologies**, the increasing **interconnectedness** amongst people in all parts of the globe and the continuous increase in the size of scientific, economic and consumer groups, has lead to a veritable **explosion in the amount of data** that is produced, used and in need of management world-wide. This is especially true in areas such as the **biological sciences**, **medical research** and **medical practice**. In these disciplines thousands of scientists, doctors and clinicians are contributing daily to a massive body of biomedical knowledge and information. Now, like never before in history, the amount of information that is available and being added to daily by the results of new scientific experiments, research and clinical trials constitutes a veritable **ocean of extraordinary depth and breadth**.

(Quotation: **Ontology for the Twenty First Century: An Introduction with Recommendations** by Andrew D. Spear)

Background

1999-2006	Diploma Thesis on "A Schema matching architecture for the bioinformatics domain" (Dipl.-Inf., University of Rostock, University of Linköping)
2009-2011	Doctorate in Database and Information Systems on "Annotation-based storage and retrieval of models and simulation descriptions in computational biology" (Dr.-Ing., University of Rostock)
2011	Post-doctoral training, Computational Neuroscience, (Norwegian University of Life Science, Ås/Norway)
Since 2012	Junior Research Group Leader "Model and Simulation Management for Systems Biology" (University of Rostock)

Outline

- What is an ontology?
- Why are ontologies so popular in biomedical research?
- Example ontologies in the life sciences & neuroscience
- Guidelines for Good Ontology Design (L. Jansen, J. Röhl)
- OWL
- Reasoning, mapping, alignment …
- Methods & Tools for ontology development: Protégé
- Practical: Modeling with Protégé (L. Schwabe)

Using ontology

Designing ontology

Managing the sea of information

The problem then, in a nutshell, is to chart the ever-growing sea of information in such a way that its various **parts**, portions and depths can be **efficiently** accessed, used, **navigated** and **reasoned** about by human individuals.

(Quotation: **Ontology for the Twenty First Century: An Introduction with Recommendations** by Andrew D. Spear)

The solution

Use computers !

- Computers reliably store huge amount of information
- they are able to efficiently, reliably and automatically retrieve and reason about the information that they store
- information stored in a computerized format can be made instantly accessible to individuals in all parts of the globe via the internet.

(Quotation: **Ontology for the Twenty First Century: An Introduction with Recommendations** by Andrew D. Spear)

Obstacle to that solution

Use computers !

- Scientists speak different languages, use different terminologies, and results are presented in different formats
- Computers… speak different languages, use different terminologies, and results are stored in different formats
- Data is not interoperable

"Data silo problem" (Barry Smith)

(Quotation: **Ontology for the Twenty First Century: An Introduction with Recommendations** by Andrew D. Spear)

Obstacle to that solution

Use computers !

“Just making data digital doesn’t make it ready for integration“ (Larson, Martone, 2009)

- How to reason with data when you find it?
- How to understand the significance of data you collected earlier?
- How to integrate with other people’s data?

(Source**: A training course in eight lectures** by Barry Smith)

So?

Use computers and ontologies!

"Define common ontologies!"

Image: http://tomgruber.org/bio/bio.htm

What is an "Ontology"?

(Quotation: **The role of common ontology in achieving sharable, reusable knowledge bases** by T. Gruber, 1991)

So?

Use computers and ontologies!

Image: http://tomgruber.org/bio/bio.htm

"Define common ontologies!"

vocabularies of representational **terms** – classes, relations, functions, object constants – with agreed-upon **definitions**, in the form of human readable text and machine-enforceable, declarative **constraints** on their **well-formed** use.
Definitions may include restrictions on **domains** and **ranges**, placement in subsumption **hierarchies**, class-wide facts **inherited** to instances, and other **axioms**.

(Quotation: **The role of common ontology in achieving sharable, reusable knowledge bases** by T. Gruber, 1991)

What is an ontology?

Origins and history

First fields of application

Philosophy: Science of Being (Aristotle, Metaphysics, IV, 1)

- What characterizes being?
- Eventually, what is being?

Artificial intelligence: existing concepts and relationships for agents

- Concentrates on what exists
- Definitions associate the names of entities in the universe of discourse (e.g., classes, relations, functions, or other objects) with human-readable text describing what the names are meant to denote, and formal axioms that constrain the interpretation and well-formed use of these terms

(Definition: T. Gruber, http://www-ksl.stanford.edu/kst/what-is-an-ontology.html)

What is an ontology?

Origins and history

Information Science

Ontologies are information artefacts that attempt to give precise formulations of the properties and relations of certain types of entities (Hofweber, 2012).

ONTOLOGIES

Better computational descriptions of science

Simon White believes that information can be better shared if everyone makes use of ontologies

***Scientific Computing World*: July/August 2005**

(Definition: GoodOD, http://www.iph.uni-rostock.de/GoodOD-Guideline.1299.0.html)

Origins and history

Computer Science

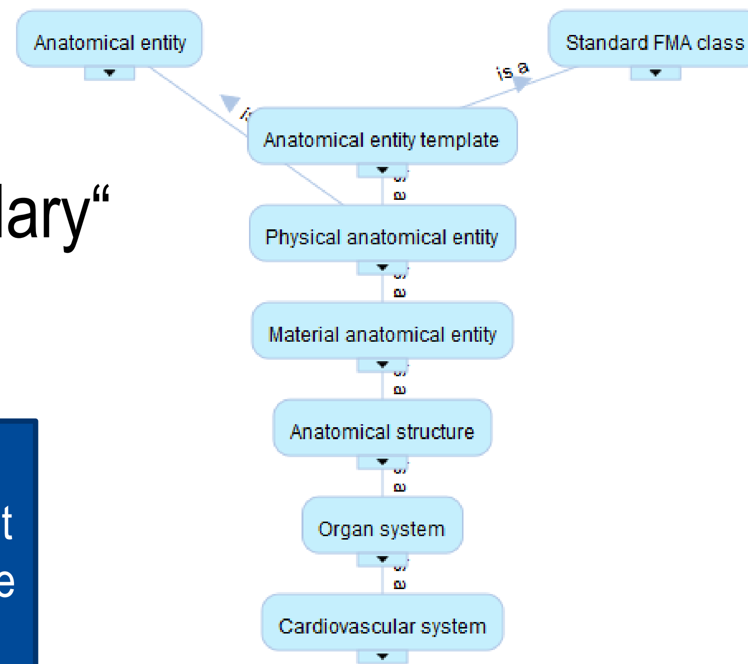
An ontology is an engineering artifact.

- Constituted by a specific vocabulary used to describe a certain reality, plus a set of explicit assumptions regarding the intended meaning of the vocabulary
- Thus, an ontology describes a formal specification of a certain domain:
    - Shared understanding of a domain of interest
    - Formal and machine manipulable model of a domain of interest
- *Definition (Gruber, 1992): „An ontology is a specification of a conceptualization.*
    - Conceptualization: couching of knowledge about the world in terms of entities
    - Specification: concrete representation of this conceptualization

(Source: http://www.co-ode.org/resources/tutorials/intro/slides/Introduction.ppt)

Defining “ontology in bio-medical research”

First approximation by Barry Smith



“a controlled, structured vocabulary“

The content of the vocabulary is kept track of by someone („curator“)

The content is structured in a way that allows reasoning on the vocabulary

(Source: **A training course in eight lectures** by Barry Smith; Figure: bioportal.bioontology.org)

What is an ontology?

Defining “ontology in bio-medical research”

Biomedical ontology

- Biomedical ontology is concerned with the principled definition of biological classes and the relations among them.
- Often associated to the idea of open data (Barry Smith)
- A prerequesite to data reuse and data linking

- Note: biomedical ontologies in practise are more than lists of terms (taxonomies), but do not necessarily meet the requirements of formal organization („ontology gradient“)

(Source: **Biomedical Ontologies** by Bodenreiter, Burgun, 2007)

Defining "ontology in bio-medical research"

Standardisation

- Standards provide
  - Common structure and terminology
  - Reducing data redundancy
- Standards allow
  - Use of common tools and techniques
  - Common training
  - Single validation

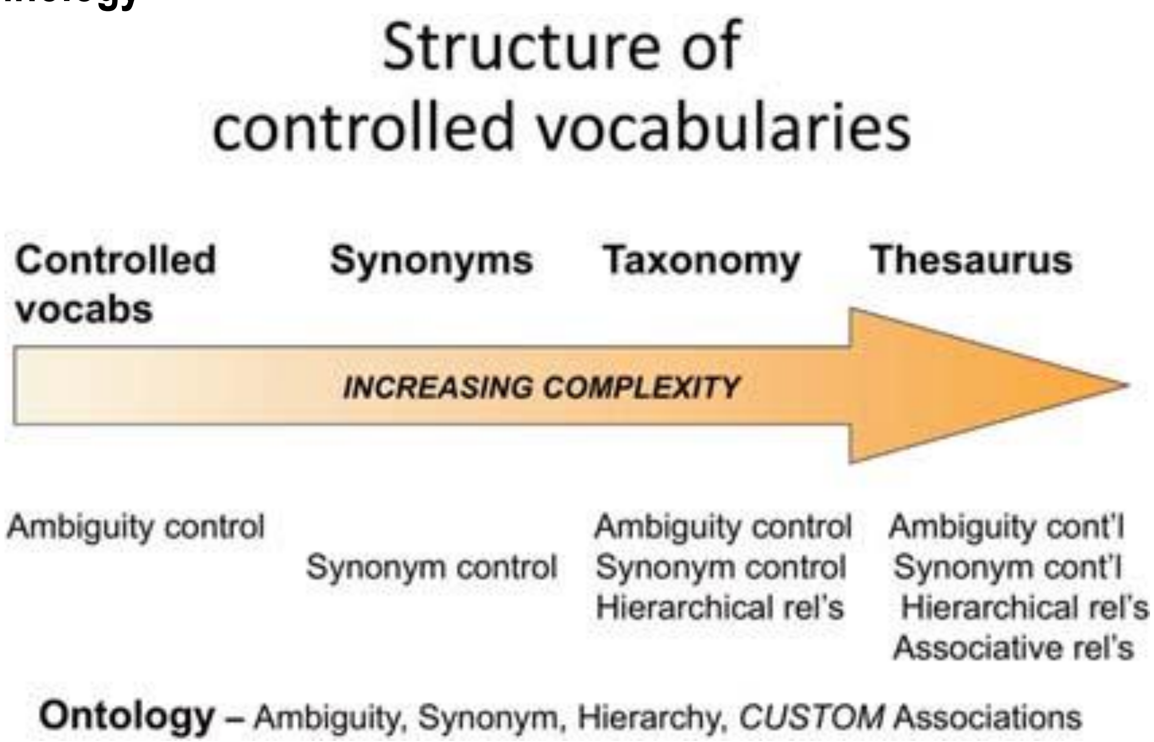
combine
the computational modeling in biology network

'Ontology' = good standards in terminology (Barry Smith, 2008)

What is an ontology?

Defining “ontology in bio-medical research”

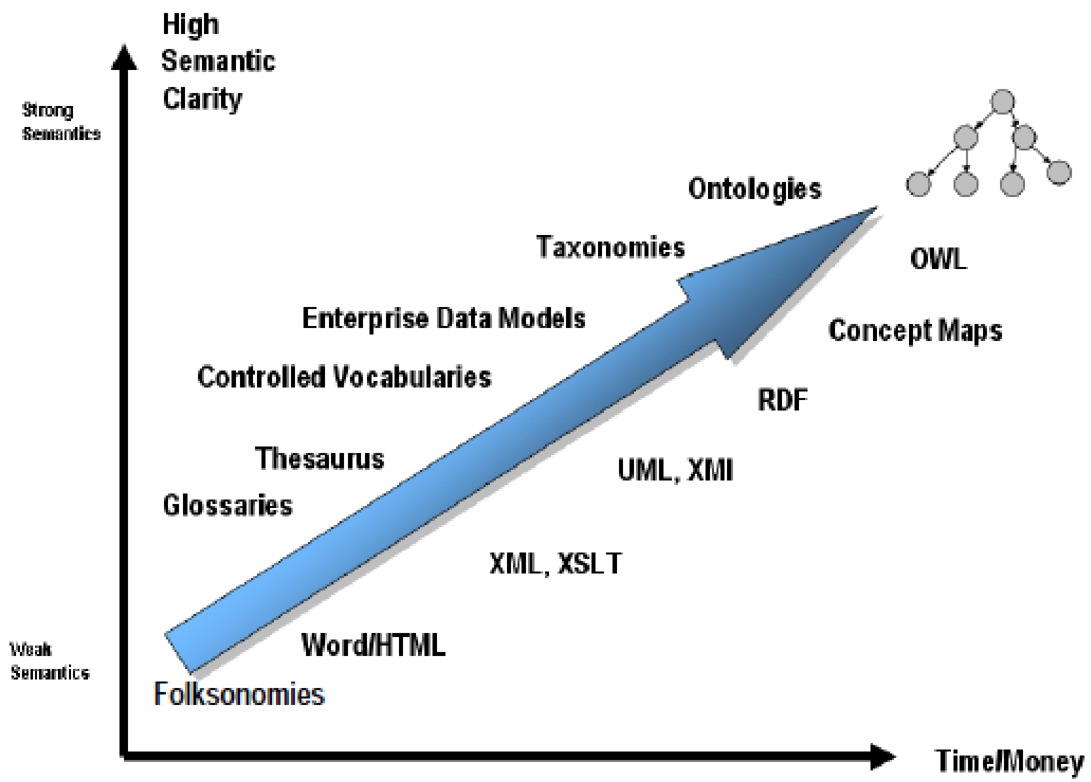
Terminology



(Source: **Building controlled vocabularies for metadata harmonization** by Zaharee, 2013)

Defining “ontology in bio-medical research”

Terminology



(Source**: A training course in eight lectures** by Barry Smith, following Leo Obrst)

Defining "ontology in bio-medical research"

Controlled vocabulary

- prescribed list of terms or headings each one having an assigned meaning (Currier et al., 2005);
- a list of terms (e.g. words, phrases) that is used to tag (label) information in a consistent way.

afferent neuron	A neuron which conveys sensory information centrally from the periphery
efferent neuron	A neuron which sends impulses peripherally to activate muscles or secretory cells.

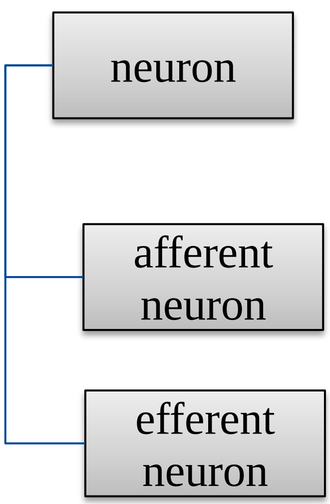
(Terms taken from the CL: purl.bioontology.org/ontology/CL)

What is an ontology?

Defining "ontology in bio-medical research"

Taxonomy

- Hierarchical classification of defined groups of things on the basis of shared characteristics



A neuron which conveys sensory information centrally from the periphery

A neuron which sends impulses peripherally to activate muscles or secretory cells.

(Terms taken from the CL: purl.bioontology.org/ontology/CL)

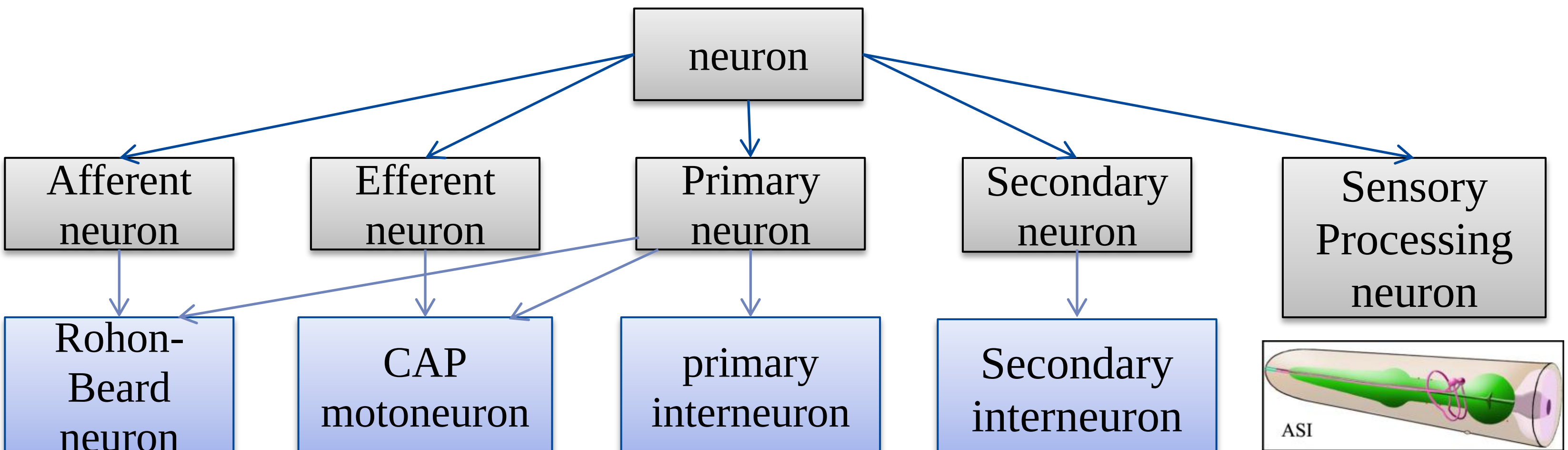
Defining “ontology in bio-medical research”

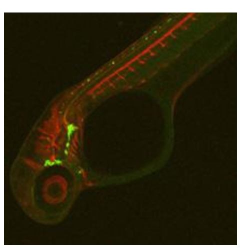
Ontology adds Relations

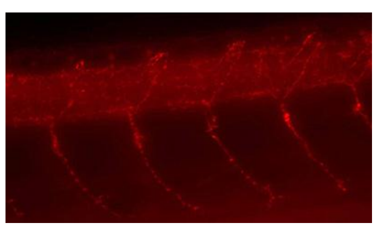
- Terms in an ontology are hierarchically structured, with the most general term on top and the most specific term at the bottom
- Terms are put into qualified relation with each other
- Example: Neuron definitions in the Cell Type ontology http://purl.bioontology.org/ontology/CL

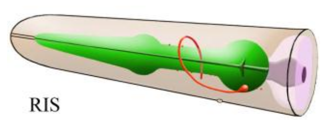
Defining “ontology in bio-medical research”

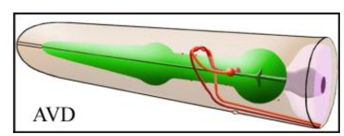
Ontology adds Relations: Neuron definitions in CL











Sources: http://thalamus.wustl.edu/nonetlab/ResearchF/elegans.html (top); http://www.biologie.ens.fr/bmddzie/spip.php?article112 & http://medschool.slu.edu/pharmphys/index.php?page=mark-m-voigt-ph-d (left)

(Terms taken from the CL: purl.bioontology.org/ontology/CL)

What is an ontology?

Defining “ontology in bio-medical research”

Requirements

…terminology, definitions, etc. that are entered into biomedical information databases should be **interoperable** between databases, internally coherent and **well-defined**, and **accurate** to the facts of reality as reflected in the current (and developing) state of **knowledge** possessed in the biomedical sciences

(Quotation: **Ontology for the Twenty First Century: An Introduction with Recommendations** by Andrew D. Spear)

Defining “ontology in bio-medical research”

Good ontologies

Formal

- The meaning of terms in an ontology is unambiguously defined to avoid misunderstanding and stated using mathematical axioms and definitions that enable automated reasoning.

Explicit

- make domain assumptions explicit for reasoning and support human understanding of a domain.

Adequate

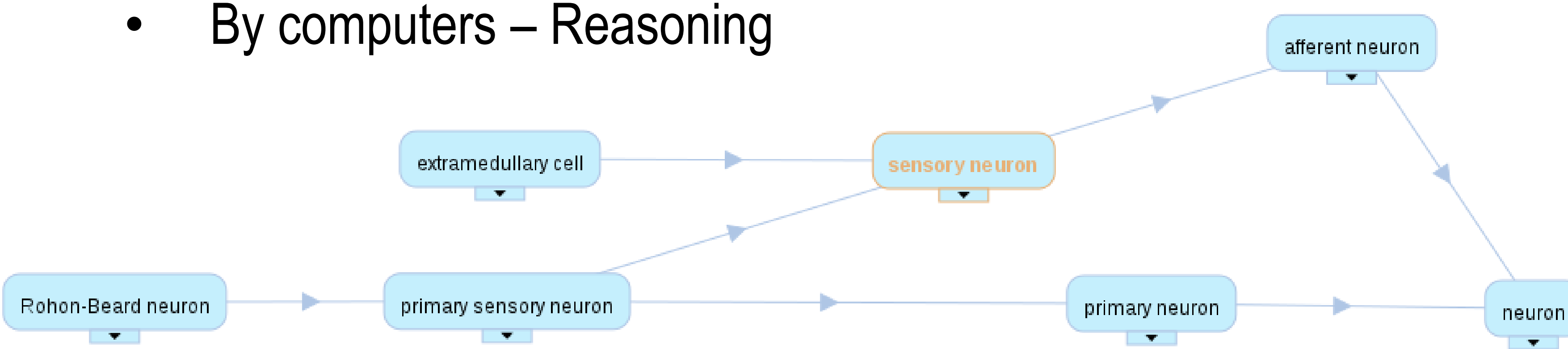
- to the domain to be represented and thus have to reflect current scientific knowledge available about the domain to be modeled.

(Definition: GoodOD, http://www.iph.uni-rostock.de/GoodOD-Guideline.1299.0.html)

Defining “ontology in bio-medical research”

Ontology information is interpretable

- By humans
- By computers – Reasoning



(Visualization: Bioportal http://bioportal.bioontology.org/ontologies/46604?p=terms&conceptid=CL%3A0000527&jump_to_nav=true#visualization)

What is an ontology?

Ontologies as starting points

Ontologies

Declare structure → **Knowledge bases Databases**

Provide domain description ↓

Software agents

Problem-solving methods

The "Semantic Web"

Open linked data

Domain-independent applications

(Figure adapted from **A Practical Introduction to Ontologies & OWL** by Alan Rector http://www.co-de.org/resources/tutorials/intro/slides/Introduction.ppt)

Ontologies in bio-medical research

(Figure: http://vimeo.com/5344028)

Reasons to use ontologies in biology

- Biology and medicine are two disciplines that rely on previously obtained knowledge
- Biology is a data rich discipline and this data needs to be made comparable, i.e. it must be integrated
- Biology has always dealt with categorization and structuring of systems, e.g. tree of life, where living organisms are classified into Kingdom, Phylum, and Class

(Source: **Ontology-based knowledge representation for bioinformatics** by Stevens et al., 2000)

Reasons to use ontologies in biology

- Community reference system (comparison of results to other project results worldwide)
- Common access to information
- Explicit relationships and underlying logic allow for automated reasoning to related entities
- Explicit bridging relationships between different ontologies for exploring underlying mechanisms

(Source: **Ontology-based knowledge representation for bioinformatics** by Stevens et al., 2000)

Applications

Use Cases: Representation of encyclopedic knowledge

- structure and make explicit encyclopedic biomedical knowledge in a form that is accessible to both researchers and machines
- “reference ontologies”, e.g. FMA specifies canonical knowledge for a domain, using a comprehensive set of entities and a large set of relationships

- Portion of neural tissue
  - Neuron
    - Axon
      - Myelinated segment of axon
      - Initial segment of axon
      - Node of Ranvier
      - Axon collateral
      - Axon terminal
    - Cell body of neuron
    - Dendrite
    - Plasma membrane of neuron
    - Protoplasm of neuron
  - Vasculature of portion of neural tissue

<u>List of attributes for neuron</u>

- Preferred name
- Synonym
- DMAID
- Definition
- Part (e.g., Axon)
- Part of (portion of neural tissue)
- Regional part (e.g., dendrite)
- Constitutional part (e.g., plasma membrane of neuron)
- Constitutional part of
- Attributed part
- Has mass / has boundary ...

Applications

Use Cases: Data Annotation

Annotation: A comment attached to a particular section of an object

- Annotation with ontologies is typically performed after (high throughput) data is acquired, often by dedicated annotators or curators
- Annotators: Process textual metadata to automatically tag text with as many ontology terms as possible, ie. combination of text mining and ontology
- E.g.: NCBO Annotator http://bioportal.bioontology.org/annotator

Use of Gene Ontology Annotation to understand the peroxisome proteome in humans

Prudence Mutowo-Meullenet*, Rachael P. Huntley, Emily C. Dimmer, Yasmin Alam-Faruque, Tony Sawford, Maria Jesus Martin, Claire O'Donovan and Rolf Apweiler

(Source: http://www.slideshare.net/drnigam/how-bio-ontologies-enable-open-science-presentation ; Larson, Martone, 2009)

Applications

Use Cases: Data Exploration

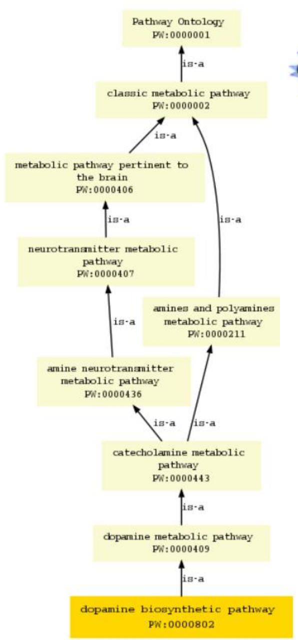
DOPAMINE BIOSYNTHETIC PATHWAY (PW:0000802)

Rat | Mouse | Human ☑ show annotations for term's descendants Sort by: symbol ↑ asc

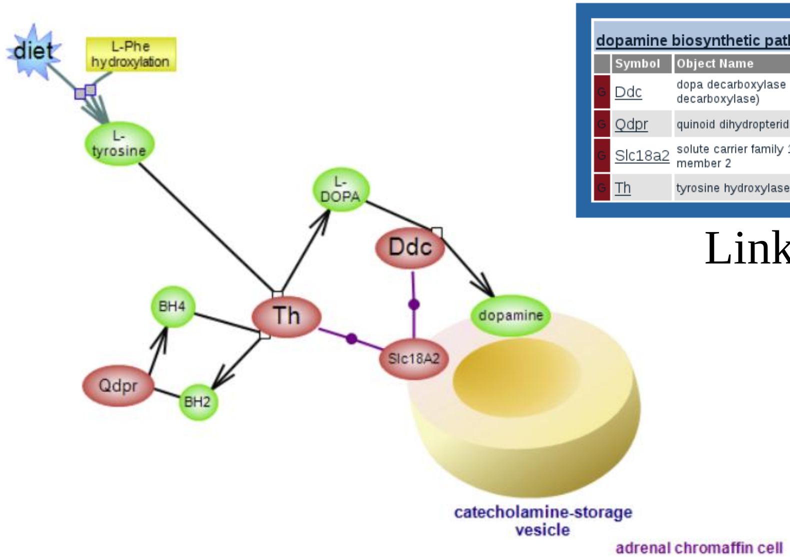
dopamine biosynthetic pathway

	Symbol	Object Name	GBrowse	Evidence	Chr	Start	Stop	Reference	Source
G	Ddc	dopa decarboxylase (aromatic L-amino acid decarboxylase)	GBrowse	ISS	14	92,698,636	92,788,635	RGD:4139904	RGD
G	Qdpr	quinoid dihydropteridine reductase	GBrowse	ISS	14	70,741,998	70,755,600	RGD:4139904	RGD
G	Slc18a2	solute carrier family 18 (vesicular monoamine), member 2	GBrowse	IDA	1	265,789,917	265,824,551	RGD:5128868	RGD
G	Th	tyrosine hydroxylase	GBrowse	ISS	1	203,164,253	203,171,506	RGD:4139904	RGD

Links to associated genes in the pathway



Classification



Diseases/Genes	Genes/Diseases

Disease Terms	Gene Symbols
Brain Injuries	Th
Brain Ischemia	Th
Contusions	Th
Depressive Disorder	Slc18a2
Diabetes Mellitus, Experimental	Qdpr , Th
Diabetes Mellitus, Type 1	Slc18a2
Dyskinesia, Drug-Induced	Ddc
Epilepsy	Th
Heart Failure	Th
Hyperinsulinism	Ddc
Hyperprolactinemia	Th
Hypertension	Th , Qdpr

Links to associated diseases and pathways

Applications

Use Cases: Data Query

- E.g., search for “tumors in mouse brain”

by allowing grouping of annotations

brain	20
hindbrain	15
rhombomere	10

Query brain without ontology 20
Query brain with ontology 45

(Source: **A training course in eight lectures** by Barry Smith, 2008)

Applications

Use Cases: Data Search of heterogeneous data

- e.g., search for "the process of creating glucose": 'glucose synthesis', 'glucose bio-synthesis', 'glucose formation', 'glucose anabolism', 'gluconeogenesis'
- An ontology can provide one class/identifier with several alternative names for the same concept (synonymy), abbreviations and acronyms

gluconeogenesis

Term information ⬇ Term neighborhood ⬇ External references ⬇ 1220 gene product associations ➡

Term Information

Accession	GO:0006094
Ontology	**Biological Process**
Synonyms	**exact**: glucose biosynthesis **exact**: glucose biosynthetic process
Definition	The formation of glucose from noncarbohydrate precursors, such as pyruvate, amino acids and glycerol. *Source:* MetaCyc:GLUCONEO-PWY
Comment	None
Subset	Prokaryotic GO subset
Community	Add usage comments for this term on the GONUTS wiki.

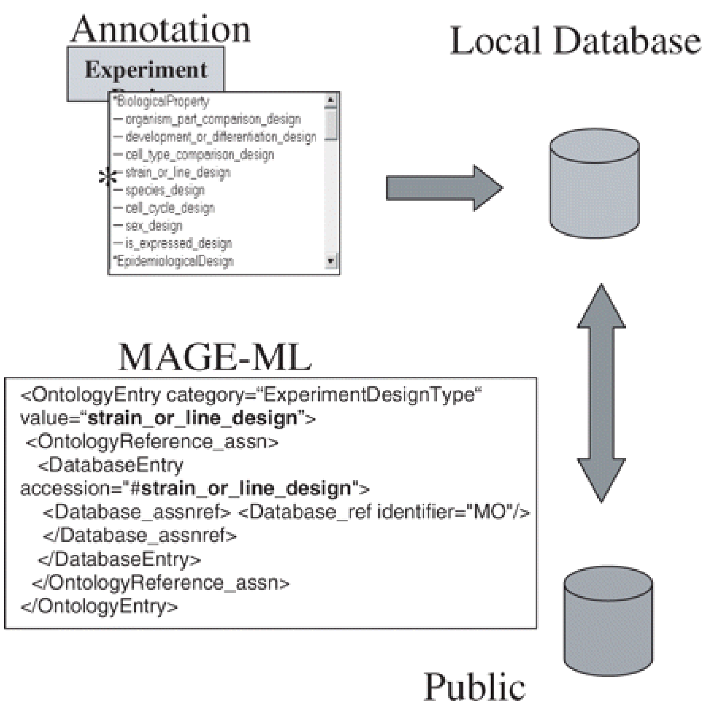
(Source: **Biomedical ontologies: a functional perspective** by .Rubin et al., 2008; Figure: Gene Ontology)

Applications

Use Cases: Data exchange among applications

Ontologies can specify how data is to be exchanged with resources

- Ontologies as explicit specification of terms used to express the biomedical information
- Relationships among data types in databases
- E.g., MGED Ontology for standardized description of microarray experiments (MO)



(Source: **The MGED Ontology: a resource for semantics-based description of microarray experiments** by Whetzel et al., 2003)

Applications

Use Cases: Reasoning with data

- Use ontologies to infer knowledge they contain
- Integrate current knowledge about a system from different data repositories
- Allow researchers to pose and test hypotheses (for consistence with the knowledge base)

Applications

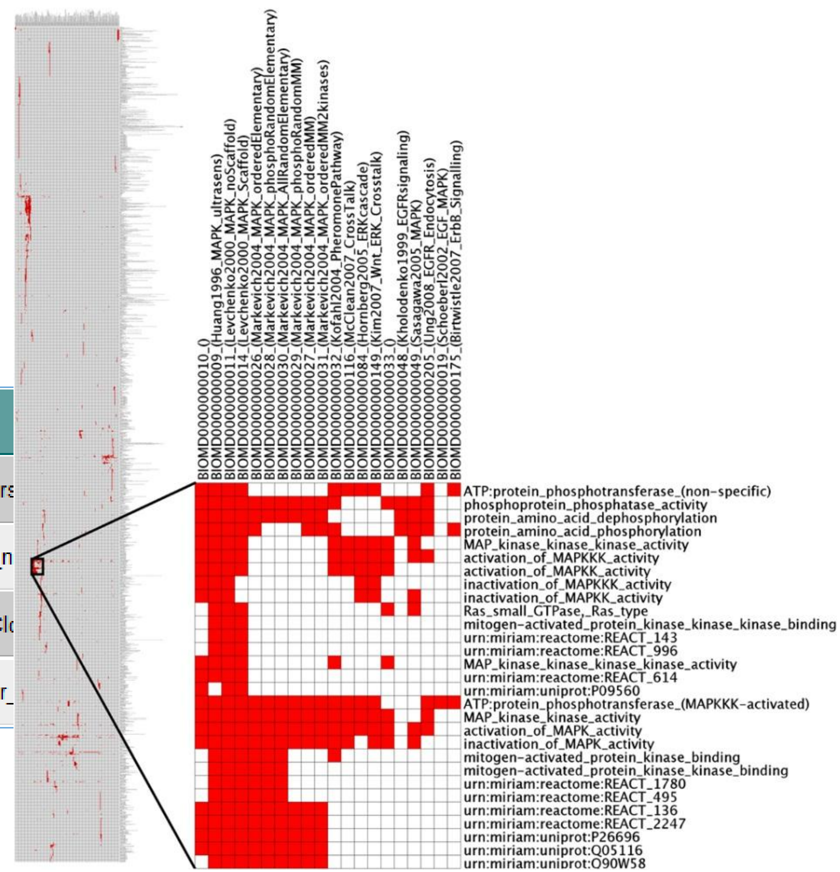
Use Cases: Model Management

- Annotation based search, retrieval and comparison

⊟ 4 Curated Models returned.

Rank	BioModels ID	Name
1. (0.0015)	BIOMD0000000118	Golomb2006_SomaticBurs
2. (0.0015)	BIOMD0000000119	Golomb2006_SomaticBursting_n
3. (0.0011)	BIOMD0000000185	Locke2008_Circadian_Cl
4. (9.1442)	BIOMD0000000060	Keizer1996_Ryanodine_receptor_

(Source: demo search in BioModels Database, http://www.ebi.ac.uk/biomodels-main/)

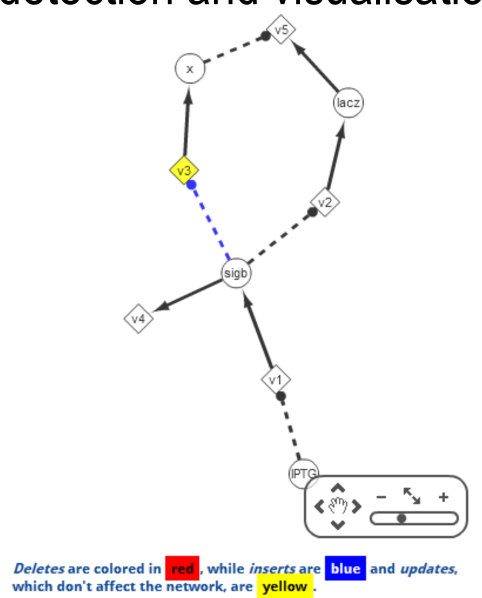


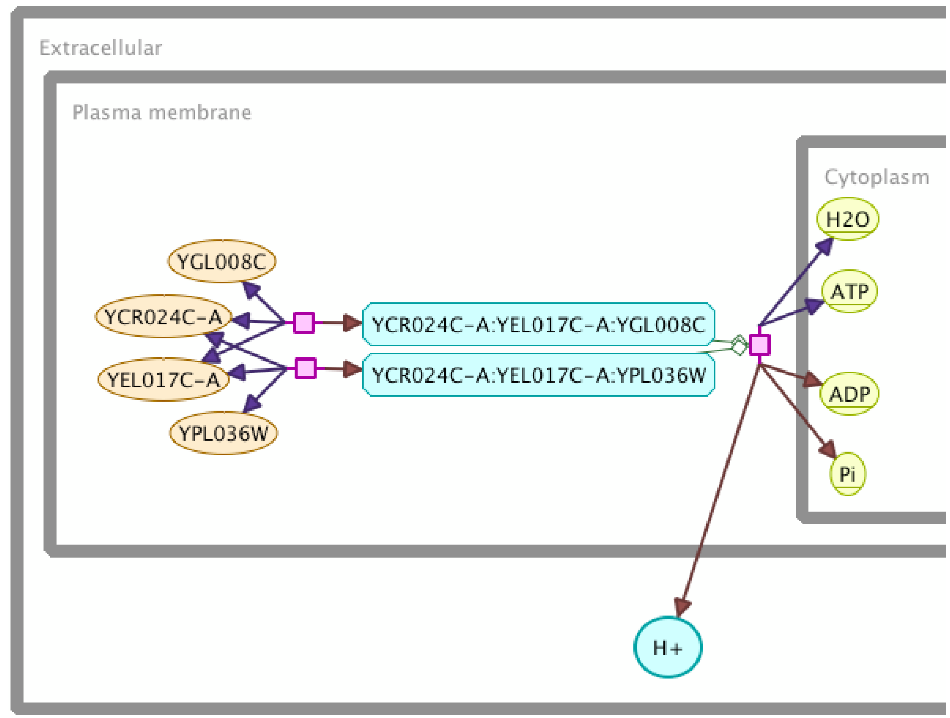
(Source: **Retrieval, alignment, and clustering of computational models based on semantic annotations** by Schulz et al., 2011)

Applications

Use Cases: Model Management

- Annotation-based difference detection and visualisation





(Source: Difference detection in Budhat, http://budhat.sems.uni-rostock.de/)

(Source: Arcadia. http://arcadiapathways.sourceforge.net/)

Applications

Use Cases: Semantic similarity in text mining


.ploscompbiol.org/article/info%3Adoi%2F10.1371%2Fjournal.pcbi.1000937

RESEARCH ARTICLE OPEN ACCESS

Semantic Similarity for Automatic Classification of Chemical Compounds

Article | Metrics | Related Content | Comments: 1

João D. Ferreira*, Francisco M. Couto
LaSIGE, University of Lisbon, Lisbon, Portugal

Abstract Top

With the increasing amount of data made available in the chemical field, there is a strong need for systems capable of comparing and classifying chemical compounds in an efficient and effective way. The best approaches existing today are based on the structure-activity relationship premise, which states that biological activity of a molecule is strongly related to its structural or physicochemical properties. This work presents a novel approach to the automatic classification of chemical compounds by integrating semantic similarity with existing structural comparison methods. Our approach was assessed based on the Matthews Correlation Coefficient for the prediction, and achieved values of 0.810 when used as a prediction of blood-brain barrier permeability, 0.694 for P-glycoprotein substrate, and 0.673 for estrogen receptor binding activity. These results expose a significant improvement over the currently existing methods, whose best

To **add a note**, highlight some text. Hide notes
Make a general comment

Jump to
Abstract
Author Summary
Introduction
Results
Discussion
Methods
Supporting Information
Author Contributions
References

View All Figures

Download: PDF
Print article
EzReprint New &

Published in the
Computational Biolo

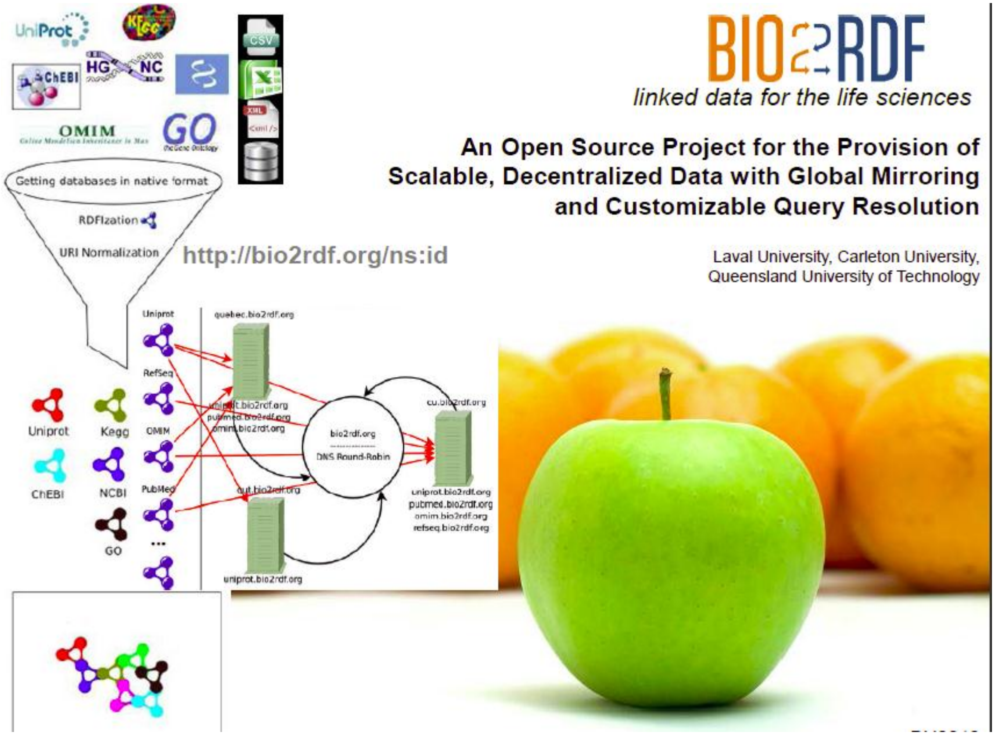
Metrics
Total Article View
Citations
CrossRef (3)
PubMed Central (3)
Scopus (2)
Web of Science® (2

Related Content
Related Articles o
Google Scholar
PubMed

Share this Article

Applications

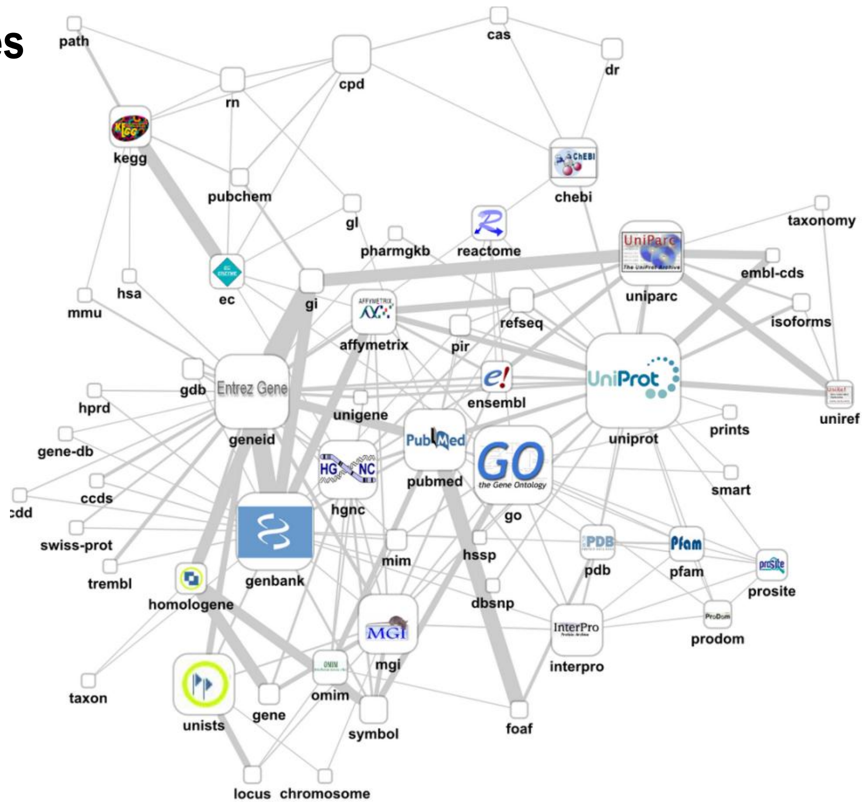
Linking data



(Fig.: **Bio2RDF** by M. Dumontier, Biohackathon 2012, Knowledge Discovery using an Integrated Semantic Web)

Applications

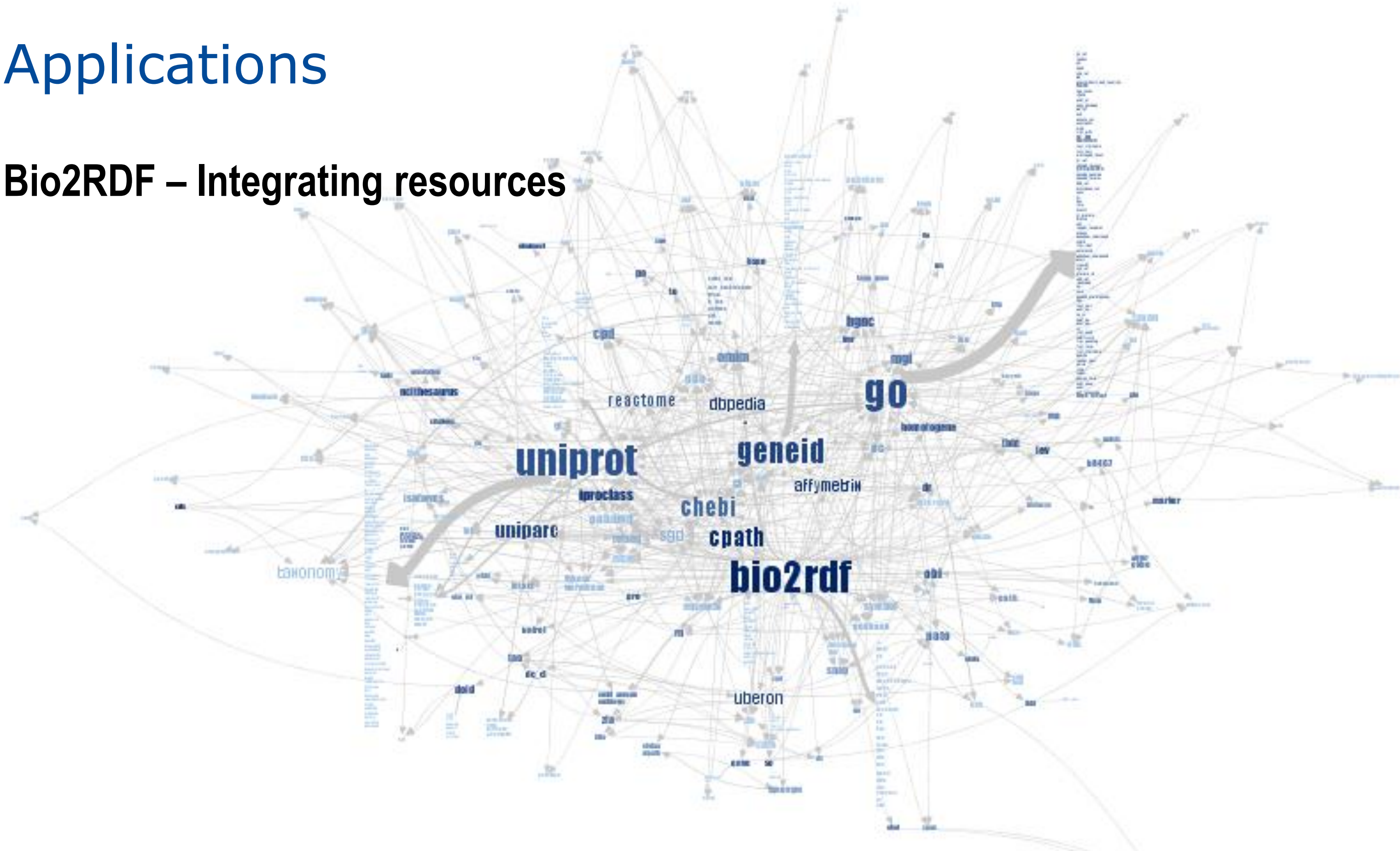
Bio2RDF – Integrating resources



(2008 @ ISMB2008 technology track)

Applications

Bio2RDF – Integrating resources



(Source: **Bio2RDF** by M. Dumontier, Biohackathon 2012, Knowledge Discovery using an Integrated Semantic Web)

Created on Many Eyes (http://many-eyes.com) © IBM

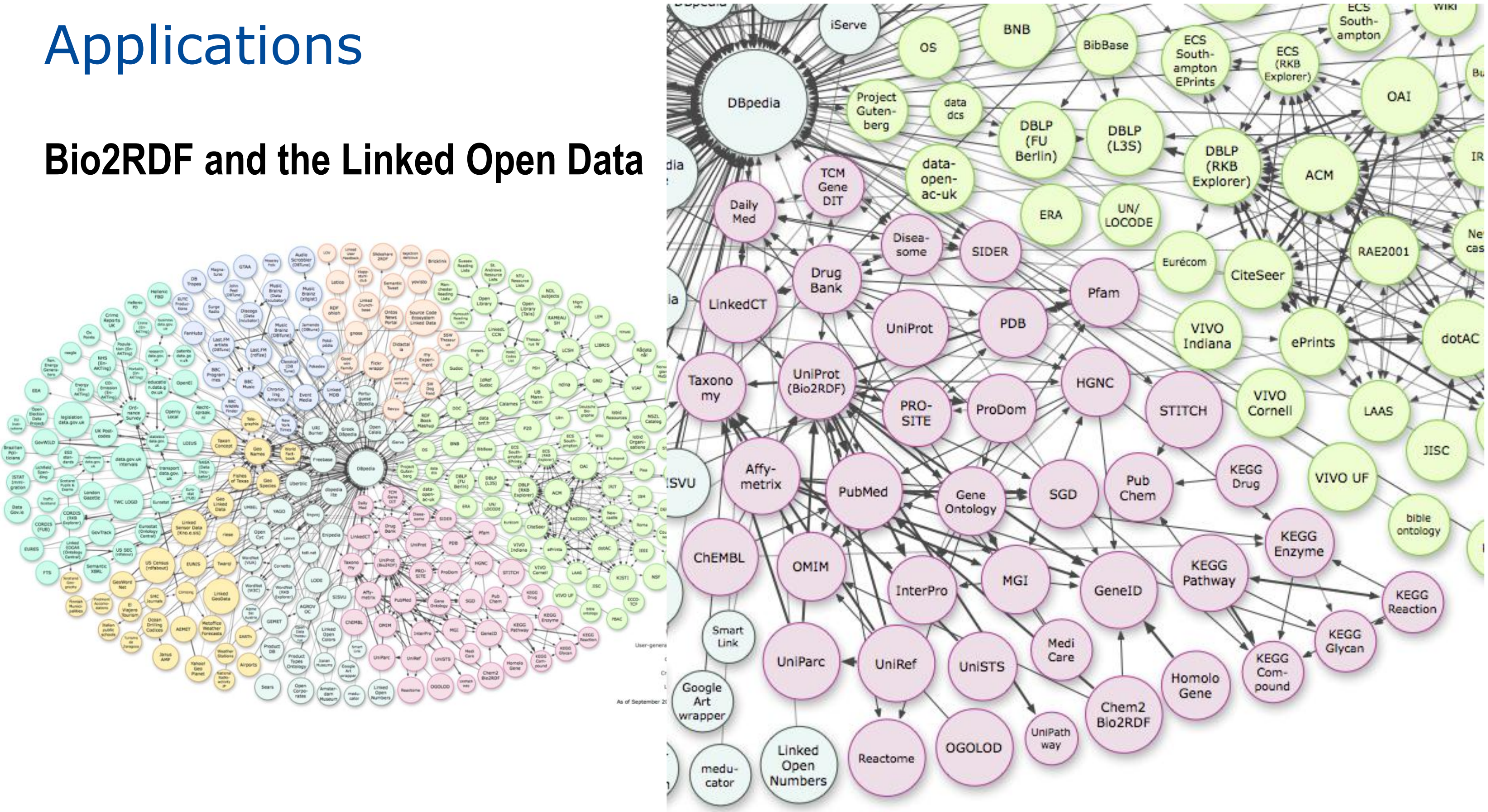
Universität Rostock Traditio et Innovatio

SEMS simulation experiment management system

Why are ontologies so popular in bio-medical research?

Applications

Bio2RDF and the Linked Open Data



(Source: LinkedOpen Data Cloud 2011, by Richard Cyganiak and Anja Jentzsch. http://lod-cloud.net/)

Applications

Bio2RDF – Integrating resources

What is Linked Open Data (LOD) and why is it good for you?

Ask Europeana: http://vimeo.com/36752317
Or Tim Berners-Lee directly:
http://www.ted.com/talks/tim_berners_lee_the_year_open_data_went_worldwide.html

Applications

How do you use ontologies in your daily work?

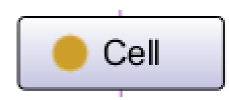
More information in the afternoon talk

Components of an ontology

„An ontology is a specification of a conceptualization.“ (Gruber, 1992)

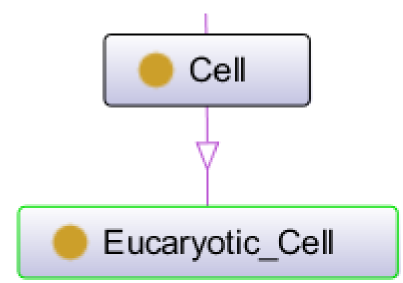
Class

- set of entities or ‘things’ within a domain; e.g.: *eukaryotic cells (are kinds of cells that have a nucleus)*
- Classes represent the canonical description of an entity



Relation

- describes the interactions between classes or a classes properties; e.g. *Eukaryotic Cell isSubclass of Cell*



(Source: **Ontology-based knowledge representation for bioinformatics** by Stevens et al., 2000)

More information in the afternoon talk

Components of an ontology

„An ontology is a specification of a conceptualization.“ (Gruber, 1992)

Property

- captures further knowledge about the relationships between classes; e.g. *Eucaryotic_Cell hasPart Nucleus*

Axiom

- constrains values for classes; e.g. *nucleic acids shorter than 20 residues are oligonucleotides*

Instance

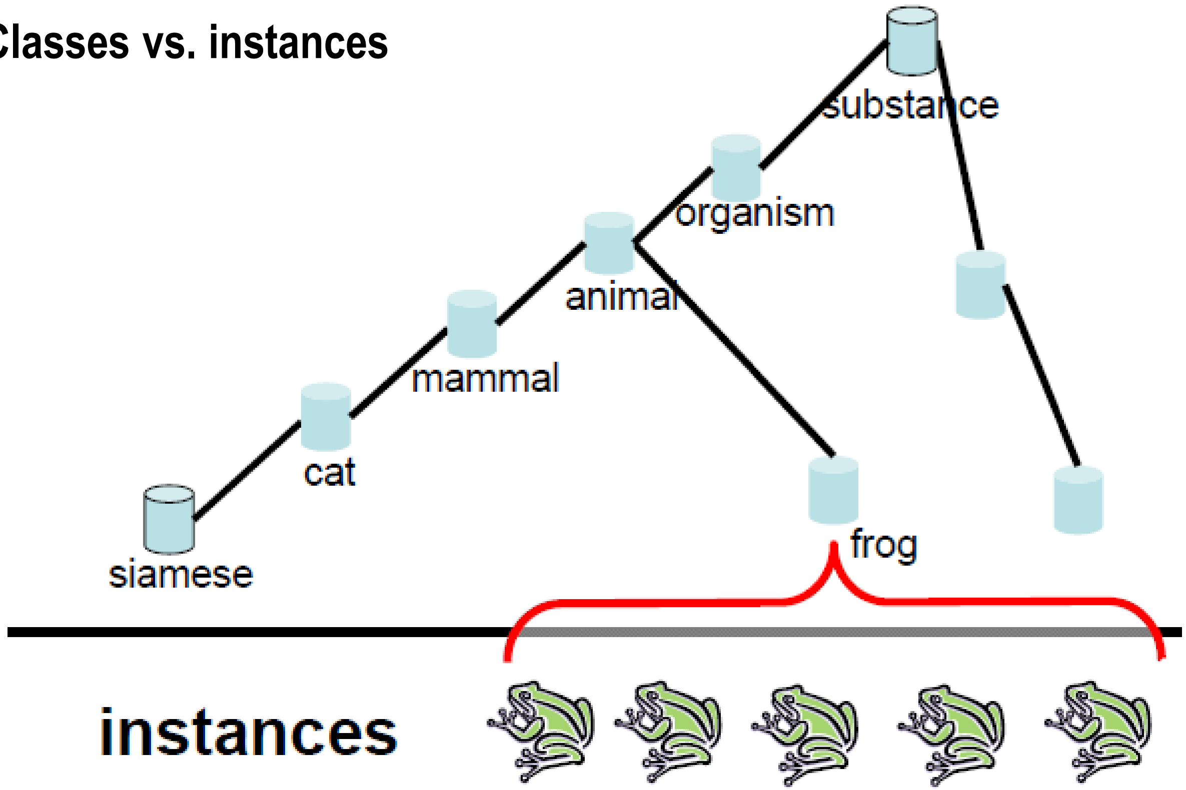
- ‘things’ represented by a concept; e.g.: *human cytochrome C is an instance of the concept Protein*
- Instances represent the individual example of a class

(Source: **Ontology-based knowledge representation for bioinformatics** by Stevens et al., 2000)

More information in the afternoon talk

Components of an ontology

Classes vs. instances



Adapted from: **A training course in eight lectures** by Barry Smith (2008)

Operations on ontologies

Merge

- Creation of a new ontology by linking existing ones
- Requirement: New ontology must contain all knowledge from the original ones

Mapping

- Expressing how the concepts in one ontology can be translated into another one
- Note: one-to-one mappings are typically not possible

Alignment

- mapping ontologies in both directions (with modification of original ontologies)

(Source: http://www.obitko.com/tutorials/ontologies-semantic-web/operations-on-ontologies.html)

Operations on ontologies

Refinement

- mapping from ontology A to ontology B so that every concept of A has equivalent in B

Unification

- Two-way refinement

Integration

- looking for the same parts of two different ontologies A and B while developing new ontology C that allows to translate between ontologies A and B

Inheritance

- ontology A inherits everything from ontology B (classes, relations, axioms) ; no inconsistency introduced by additional knowledge contained in ontology (Modularity)

(Source: http://www.obitko.com/tutorials/ontologies-semantic-web/operations-on-ontologies.html)

Components of an ontology

„An ontology is a specification of a conceptualization." (Gruber, 1992)

The components of an ontology need to be described, stored and exchanged in a formal manner.

Data exchange standards for ontologies

Languages for explicit specification

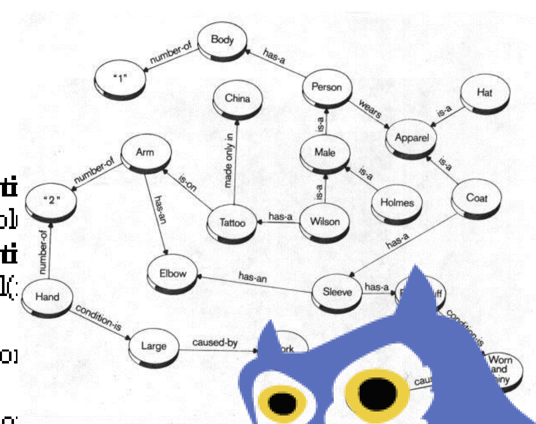
- In order to manage and perform computations on the different relationships between entities, ontologies are usually encoded into a language that allows a machine to manage and utilize the information.
- Notion of an ontology is independent of the language in which it is encoded

Data exchange standards for ontologies

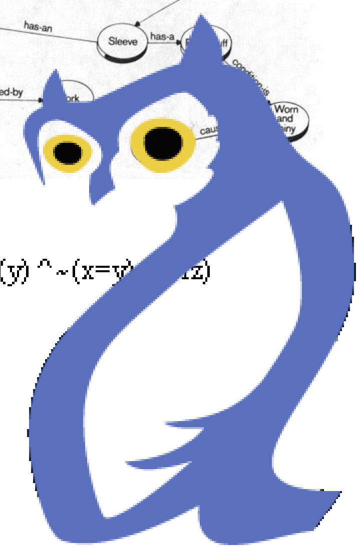
More information in the afternoon talk and tomorrow

Languages for explicit specification

- Graphical
  - Semantic networks
  - topic maps
  - UML
- Logic-based
  - Description logics
  - Rules
  - First order logic
  - Conceptual graphs



Every gardener likes the sun.
(Ax) gardener(x) => likes(x,Sun)
You can fool some of the people all of the ti
(Ex)(At) (person(x) ^ time(t)) => can-fool
You can fool all of the people some of the ti
(Ax)(Et) (person(x) ^ time(t) => can-fool(
All purple mushrooms are poisonous.
(Ax) (mushroom(x) ^ purple(x)) => poiso
No purple mushroom is poisonous.
~(Ex) purple(x) ^ mushroom(x) ^ poisono
(Ax) (mushroom(x) ^ purple(x)) => ~poisonous(x)
There are exactly two purple mushrooms.
(Ex)(Ey) mushroom(x) ^ purple(x) ^ mushroom(y) ^ purple(y) ^ ~(x=y) ... z)
(mushroom(z) ^ purple(z)) => ((x=z) v (y=z))



Class: NuclearMembraneReceptorGeneProduct

EquivalentTo: GeneProduct

that *has_molecular_function* some ReceptorActivity

(Sources: http://www.cs.man.ac.uk/%7Ehorrocks/Teaching/cs646/Slides/ontologies.ppt; **Exploring Gene Ontology Annotations with OWL** by Jupp et al., 2011)

More information in tomorrow‘s talk

Data exchange standards for ontologies

Web Ontology Language (OWL)

- W3C standard with a large and growing ecosystem of developers
- Description Logics
- Using OWL in Protégé you can:
  - Query your ontology
  - E.g., for sophisticated queries on your website
  - Quickly find mistakes
  - Automate classification
- “Manchester syntax” for more intuitive ontology modeling
- Note: Round tripping from OBO to OWL allows to develop in OBO while taking advantage of OWL and Protégé for reasoning

(Sources: http://www.cs.man.ac.uk/%7Ehorrocks/Teaching/cs646/Slides/ontologies.ppt; **Exploring Gene Ontology Annotations with OWL** by Jupp et al., 2011)

More information in tomorrow's talk

Data exchange standards for ontologies

Open Biomedical Ontologies (OBO) format

- controlled vocabularies for shared use across different biological and medical domains
- subset of the concepts in the OWL description logic language, with several extensions for meta-data modeling
- text file format for OBO ontologies
- Goals:
  - Human readability
  - Ease of parsing
  - Extensibility
  - Minimal redundancy

Universität Rostock Traditio et Innovatio

SEMS simulation experiment management system

Summary Part 1

"An ontology may take a variety of forms, but necessarily it will include a vocabulary of terms, and some specification of their meaning. This includes definitions and an indication of how concepts are inter-related which collectively impose a structure on the domain and constrain the possible interpretations of terms."

(Quotation: **The enterprise ontology** by Uschold et al., 1998)

Universität Rostock Traditio et Innovatio

SEMS simulation experiment management system

References Part 1

Bodenreiter O, Burgun A (2007) *Biomedical Ontology.* Book Chapter, http://ai.arizona.edu/mis596a/book_chapters/medinfo/Chapter_08.pdf (accessed 2013-02-06)

Dumontier, M (2012) *Knowledge Discovery using an Integrated Semantic Web*. Biohackathon.

Gruber, TR (1991) *The role of common ontology in achieving sharable, reusable knowledge bases.* In: Proceedings of the Second International Conference on Principles of Knowledge Representation and Reasoning.

Jupp S, Stevens R, Hoehndorf R (2011) *Exploring Gene Ontology Annotations with OWL.* http://www.cs.man.ac.uk/~sjupp/mouse_GOA/Jupp-bio-ontologies-2011-final.pdf (Accessed: 2013-02-06)

Larson SD and Martone ME (2009) *Ontologies for neuroscience: what are they and what are they good for?* Frontiers in neuroscience 3(1):60

Rector, A (2005) A Practical Introduction to Ontologies & OWL, Manchester University, **http://www.co-ode.org/resources/tutorials/intro/slides/Introduction.ppt** (accessed 2013-01-25)

Schulz M, Krause F, Le Novère N, Klipp E, Liebermeister W (2011) *Retrieval, alignment, and clustering of computational models based on semantic annotations*. Molecular systems biology 7(1)

Smith, B *A training course in eight lectures*, http://ontology.buffalo.edu/smith/BioOntology_Course.html (accessed 2013-02-06)

Spear, AD (2005) *Ontology for the Twenty First Century: An Introduction with Recommendations.* Internal report

Stevens R, Goble CA , Bechhofer S (2000) *Ontology-based knowledge representation for bioinformatics*. Briefings in Bioinformatics 1(4): 398-414

References Part 1 (cont'd)

Uschold M, King M, Moralee S, Zorgios Y (1998): *The enterprise ontology*. Knowledge Eng. Rev., Vol. 13(1) (special issue on 'Putting ontologies to use'), pp. 31–89.

Whetzel PL, Parkinson H, Causton HC et al. (2003) The MGED Ontology: a resource for semantics-based description of microarray experiments. Bioinformatics 22(7)

Zaharee, M (2013) *Building controlled vocabularies for metadata harmonization.* Bulletin of the American Society for Information Science and Technology Volume 39, Issue 2, pages 39-42, 14 DEC 2012 DOI: 10.1002/bult.2013.1720390211

http://onlinelibrary.wiley.com/doi/10.1002/bult.2013.1720390211/full#fig1