

# Semantic Web Technologies for Biomedicine

---

Natasha Noy  
Stanford Center for Biomedical Informatics Research  
Stanford University

# Why Medicine and Biology?

---

- **LOTS** of data
- The data is in different database silos
- Considerable benefit if you can integrate information

AATDB, AceDb, ACUTS, ADB, AFDB, AGIS, AMSdb, ARR, AsDb, BBDB, BCGD, Beanref, Biolmage, BioMagResBank, BIOMDB, BLOCKS, BovGBASE, BOVMAP, BSORF, BTKbase, CANSITE, CarbBank, CARBHYD, CATH, CAZY, CCD, CD4OLbase, CGAP, ChickGBASE, Colibri, COPE, CottonDB, CSNDB, CUTG, CyanoBase, dbCFC, dbEST, dbSTS, DDBJ, DGP, DictyDb, Picty\_cDB, DIP, DOGS, DOMO, DPD, DPInteract, ECDC, ECGC, EC02DBASE, EcoCyc, EcoGene, EMBL, EMD db, ENZYME, EPD, EpoDB, ESTHER, FlyBase, FlyView, GCRDB, GDB, GENATLAS, Genbank, GeneCards, Genline, GenLink, GENOTK, GenProtEC, GIFTS, GPCRDB, GRAP, GRBase, gRNAsdb, GRR, GSDB, HAEMB, HAMSTERS, HEART-2DPAGE, HEXAdb, HGMD, HIDB, HIDC, HIVdb, HotMolecBase, HOVERGEN, HPDB, HSC-2DPAGE, ICN, ICTVDB, IL2RGbase, IMGT, Kabat, KDNA, KEGG, Klotho, LGIC, MAD, MaizeDb, MDB, Medline, Mendel, MEROPS, MGDB, MGI, MHCPEP5, Micado, MitoDat, MITOMAP, MJDB, MmtDB, Mol-R-U's, MPDB, MRR, MutBase, MycDB, NDB, NRSub, O-lycBase, OMIA, OMIM, OPD, ORDB, OWL, PAHdb, PatBase, PDB, PDD, Pfam, PhosphoBase, PigBASE, PIR, PKR, PMD, PPDB, PRESAGE, PRINTS, ProDom, Prolysis, PROSITE, PROTOMAP, RatMAP, RDP, REBASE, RGP, SBASE, SCOP, SeqAnaiRef, SGD, SGP, SheepMap, Soybase, SPAD, SRNA db, SRPDB, STACK, StyGene, Sub2D, Subtilist, SWISS-2DPAGE, SWISS-3DIMAGE, SWISS-MODEL Repository, SWISS-PROT, TelDB, TGN, tmRDB, TOPS, TRANSFAC, TRR, UniGene, URNADB, VBASE, VDRR, VectorDB, WDCM, WIT, WormPep, YEPD, YPD, YPM, etc ..... !!!!

# Outline

---

- Biomedical ontologies
- Applications of Semantic Web technologies in biomedicine
- Organizations and consortia promoting Semantic Web technologies in biomedicine
- Challenges and research issues

# Biomedical Ontologies


---

- The biomedical community has embraced ontologies perhaps more than other fields
- Several **large prominent** biomedical ontologies
  - National Cancer Institute's Thesaurus (~**80K** classes, OWL)
  - Foundational Model of Anatomy (~**70K** classes, 2 million relations, Protege frames, OWL version available)
  - Gene Ontology (~**20K** classes, OBO)
- National Center for Biomedical Ontology (NCBO) BioPortal indexes **more than 200 ontologies**



# BioPortal: A place to find biomedical ontologies

<http://biportal.bioontology.org>


 BioPortal [Browse](#) [Search](#) [Projects](#) [Annotate](#) [All Mappings](#) [All Resources Alpha](#) [Sign In](#) [Register](#)

**SUBMIT ONTOLOGY** [Submit New Ontology](#)


**FILTER BY CATEGORY**

**FILTER BY GROUP**  [Link To This Filter](#)

**FILTER BY TEXT**

 [Subscribe to all ontologies](#)

| ONTOLOGY NAME   | FOR      |
|---|----------|
| <a href="#">ABA Adult Mouse Brain (ABA)</a>                                       | OWL      |
| <a href="#">African Traditional Medicine (ATMO)</a>                               | OBC Form |
| <a href="#">Amino Acid (amino-acid)</a>   | OWL      |
| <a href="#">Amino Acid with Simplified Chinese annotations (Amino Acid-zh_CN)</a> | OWL      |
| <a href="#">Amphibian gross anatomy (AAO)</a>                                     | OBC Form |
| <a href="#">Animal natural history and life history (ADW)</a>                     | PRO      |
| <a href="#">Ascomycete phenotype ontology (APO)</a>                               | OBC Form |
| <a href="#">Basic Vertebrate Anatomy (basic-vertebrate-gross-anatomy)</a>         | OWL      |
| <a href="#">Bilateria anatomy (BILA)</a>  | OBC Form |

**Biomedical Resource Ontology**  [Subscribe](#)

[Metadata](#) [Projects \(1\)](#) [Reviews \(0\)](#) [Ontology Widgets](#)

|                       |  |                            |   |
|-----------------------|--|----------------------------|---|
| <b>ONTOLOGY NAME:</b> | Biomedical Resource Ontology   | <b>CONTACT(S):</b>         | Csongor Nyulas, Natasha Noy   |
| <b>ONTOLOGY ID:</b>   | 1104   | <b>HOME PAGE:</b>          | <a href="http://www.ncbcs.org/biositem">http://www.ncbcs.org/biositem</a> |
| <b>FORMAT:</b>        | OWL  | <b>DOCUMENTATION PAGE:</b> | <a href="http://www.ncbcs.org/biositem">http://www.ncbcs.org/biositem</a> |
| <b>CATEGORIES:</b>    | Biomedical Resources   | <b>PUBLICATIONS PAGE:</b>  |   |
| <b>CONTACT EMAIL:</b> | csongor.nyulas@stanford.edu  |                            |   |
| <b>DESCRIPTION:</b>   | A controlled terminology for the 'resource_type' and which is used to improve the sensitivity and specificity of web searches. |                            |   |

**Notes**

Atlas BRO:Bibliographic\_Resource BRO:Binary\_Executable BRO:Clinical\_Data BRO:Data\_dimensionality BRO:Data\_Distribution BRO:Data\_Repository BRO:Data\_Service BRO:Funding\_Source BRO:Image BRO:Object\_Database BRO:Online\_Course BRO:Portals BRO:Research\_Supplies BRO:Resource BRO:Service BRO:Software BRO:Software\_Distribution BRO:Support BRO:Text\_Mining Data\_Distribution Document Education Identity\_Management Individual\_Human\_Data Paper Research\_Supplies\_and\_Services Software\_Distribution Software\_Documentation

**Mappings**

BRO:CT\_Scanner BRO:Electron\_Microscope BRO:Light\_Microscope BRO:Mass\_Spectrometer BRO:Microscope Software

| VERSION NUMBER | RELEASE DATE | ONTOLOGY FILE                     | VISUALIZE               | DOWNLOAD DIFF FILE  |
|----------------|--------------|-----------------------------------|-------------------------|---|
| 2.7            | 05/14/2009   | <a href="#">Download Ontology</a> | <a href="#">Explore</a> |   |
| 2.6            | 03/02/2009   | <a href="#">Download Ontology</a> | <a href="#">Explore</a> |   |
| 2.5            | 12/09/2008   | <a href="#">Download Ontology</a> | <a href="#">Explore</a> | Download Diff with previous version ( <a href="#">Txt</a>   <a href="#">RDF</a> ) |

# BioPortal features

---

- Browse and visualize (biomedical) ontologies
- Search within and across the ontologies
- View details of individual classes
- Request changes and additions from ontology authors
- Create and browse ontology mappings

# What will you find in a repository of biomedical ontologies?

---

Use of Semantic Web Technologies

- Ontologies that were developed **from the start using Semantic Web technologies**
- Ontologies that were first developed in **non-standard languages, but now use OWL**
- Ontologies in the **OBO format** and other similar formats(not OWL)
  - simple *frame-based* format
  - defines *subclass-superclass* hierarchy
  - defines *part-of* relationships
  - defines properties for *preferred names, synonyms*
- Medical **terminologies** that are now becoming ontologies

# Translational Medicine Ontology (TMO)

---

- Developed by the W3C Healthcare and Life Sciences Interest Group
- Developed as an OWL ontology
- Uses extensive user-based requirements collection
- Includes requirements from many types of users:
  - clinicians
  - informaticians
  - chemists
  - researchers
  - sales and marketing
  - ....



# National Cancer Institute's Thesaurus (NCI Thesaurus)

NCI Thesaurus Version 10.03 Melanoma | [Link Here](#) | [Subscribe](#)

View Ontology Summary Details Visualization Notes (2) Mappings (4) Resource Index

Jump To:

Legend

- Abnormal Cell
- Activity
- Anatomic Structure, System, or Substance
- Biochemical Pathway
- Biological Process
- Chemotherapy Regimen or Agent Combination
- Conceptual Entity
- Diagnostic or Prognostic Factor
- Diagnostic, Therapeutic, or Research Equipment
- Disease, Disorder or Finding
  - Disease or Disorder
    - Behavior-Related Disorder
    - Cancer-Related Condition
    - Disorder by Site
    - Genetic Disorder
    - Hamartoma
    - Hyperplasia
    - Neoplasm
      - Neoplasm by Morphology
      - Neoplasm by Site
      - Neoplasm by Special Category
        - Benign Neoplasm
        - Childhood Neoplasm
        - Common Neoplasm
          - Astrocytic Tumor
          - Common Carcinoma
          - Common Connective and
          - Common Germ Cell Tumo
          - Common Hematopoietic N
          - Melanocytic Skin Neoplas
          - Melanoma**

|                      |   |
|----------------------|---|
| ID:                  | Melanoma  |
| Full Id:             | <a href="http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#Melanoma">http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#Melanoma</a>   |
| Synonyms:            | Malignant Melanoma  |
| Definitions:         | A malignant, usually aggressive tumor composed of atypical, neoplastic melanocytes. Most often, melanomas arise in the skin (cutaneous melanomas) and include the following histologic subtypes: superficial spreading melanoma, nodular melanoma, acral lentiginous melanoma, and lentigo maligna melanoma. Cutaneous melanomas may arise from acquired or congenital melanocytic or dysplastic nevi. Melanomas may also arise in other anatomic sites including the gastrointestinal system, eye, urinary tract, and reproductive system. Melanomas frequently metastasize to lymph nodes, liver, lungs, and brain. |
| Alt Definition:      | A form of skin cancer that arises in melanocytes, the cells that produce pigment. Melanoma usually begins in a mole.NCI-GLOSS   |
| Neoplastic Status:   | Malignant   |
| Code:                | C3224   |
| Umls Cui:            | C0025202  |
| Label:               | Melanoma  |
| Equivalent Class:    | 'Common Neoplasm' and 'Melanocytic Neoplasm' and (Disease_Has_Abnormal_Cell only 'Melanoma Cell')   |
| Legacy Concept Name: | Melanoma  |
| Icd O 3 Code:        | 8720/3  |
| Semantic Type:       | Neoplastic Process  |
| Preferred Name:      | Melanoma  |
| Maps To Lash:        | Malignant Melanoma  |

# The Gene Ontology (GO): provide consistent descriptions of gene products



The screenshot shows the AmiGO website interface. At the top left is the 'AmiGO' logo. To the right is a search bar labeled 'Search GO:' with a text input field. Below the search bar are two radio buttons: 'Terms' (selected) and 'Gene Products'. Below the search bar is a navigation menu with links: 'Top Docs', 'Gene Ontology', 'GO Links', and 'GO Summary'. The main content area displays a hierarchical tree of Gene Ontology terms. The root term is 'GO:0003673 : Gene Ontology (92932)'. It is expanded to show 'GO:0008150 : biological process (56952)'. This term is further expanded to show several sub-terms, including 'GO:0007610 : behavior (566)', 'GO:0000004 : biological process unknown (6152)', 'GO:0007154 : cell communication (11916)', 'GO:0007155 : cell adhesion (830)', 'GO:0030260 : cell invasion (0)', 'GO:0008037 : cell recognition (210)', 'GO:0007267 : cell-cell signaling (1318)', 'GO:0045168 : cell-cell signaling involved in cell fate commitment (0)', 'GO:0030072 : peptide hormone secretion (6)', 'GO:0030252 : growth hormone secretion (2)', 'GO:0030073 : insulin secretion (4)', 'GO:0030103 : vasopressin secretion (2)', 'GO:0019226 : transmission of nerve impulse (688)', and 'GO:0030383 : host-pathogen interaction (12)'. Each term is preceded by a square icon with a plus sign, indicating it is expanded. The terms are color-coded with small icons to the right of the term name.

AmiGO

Search GO:

Terms  Gene Products

[Top Docs](#) [Gene Ontology](#) [GO Links](#) [GO Summary](#)

[-] **GO:0003673 : Gene Ontology (92932)**

- [-] **GO:0008150 : biological process (56952)**
  - [+] **GO:0007610 : behavior (566)**
    - **GO:0000004 : biological process unknown (6152)**
  - [-] **GO:0007154 : cell communication (11916)**
    - [+] **GO:0007155 : cell adhesion (830)**
      - **GO:0030260 : cell invasion (0)**
    - [+] **GO:0008037 : cell recognition (210)**
    - [-] **GO:0007267 : cell-cell signaling (1318)**
      - [+] **GO:0045168 : cell-cell signaling involved in cell fate commitment (0)**
      - [-] **GO:0030072 : peptide hormone secretion (6)**
        - **GO:0030252 : growth hormone secretion (2)**
        - **GO:0030073 : insulin secretion (4)**
        - **GO:0030103 : vasopressin secretion (2)**
      - [+] **GO:0019226 : transmission of nerve impulse (688)**
    - [+] **GO:0030383 : host-pathogen interaction (12)**

# Medical Terminologies “growing up” to become ontologies

---

- **International Classification of Diseases (ICD)** from the World Health Organization (WHO)
  - Used by *all* United Nations member countries to collect epidemiological and healthcare-related information
  - In its 10th edition; more than 100 years old (US uses ICD-9)
  - The first ten editions are essentially lists of terms, distributed in hard-bound volumes
- **ICD-11 will be a formal OWL ontology**
  - developed collaboratively by members of the international community
  - using iCAT, a custom-tailored version of WebProtégé (<http://icatdemo.stanford.edu>)



# iCAT: An adaptation of Protege for ICD-11

**iCAT DEMO and TRAINING PLATFORM** You are signed out. | [Sign In](#) | [Send feedback](#)

My ICD | **ICD Content** | Category Notes and Discussions | Reviews | Change History | Manage Hierarchy

### ICD Categories

Create Watch

- ICD Categories (1) (4320)
- Superparents (1) (1)
- 01 I Certain infectious and parasitic diseases (443)
- 02 II Neoplasms (305)
- 03 III Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism
- 04 IV Endocrine, nutritional and metabolic diseases (93)
- 05 V Mental and behavioural disorders (68)
- 06 VI Diseases of the nervous system (484)
- 07 VII Diseases of the eye and adnexa (2) (404)
- 08 VIII Diseases of the ear and mastoid process (2)
- 09 IX Diseases of the circulatory system (22)
- 10 X Diseases of the respiratory system (25)
- 11 XI Diseases of the digestive system (334)
- 12 XII Diseases of the skin (5) (1607)
- LA Infections and infestations affecting the skin (2) (420)**
  - LA0 Viral infections affecting the skin (2) (83)
  - LA1 Bacterial infections affecting the skin (1) (186)
  - LA2 Fungal infections affecting the skin (1) (95)
  - LA4 Parasitic infestations affecting the skin (1) (33)
  - LA5 Infestation of the skin by arthropods (1) (14)
  - LAz Infection of skin - other and unspecified (1) (2)
  - LB Inflammatory dermatoses (1) (212)

### Details for LA Infections and infestations affecting the skin

Title & Definition | Classification Properties | Terms | **Body Structure Description** | Manifestation Properties

Causal Properties | Temporal Properties | Severity Properties | Functional Properties | Specific Condition Properties

Treatment | Diagnostic Criteria | ICD 10 Notes and Hints

#### Body System

| label                        | Term ID | L... |   |   |
|------------------------------|---------|------|---|---|
| Skin and subcutaneous tissue |         |      | X | 🗨 |

+ Find term  
+ Add term

#### Body Part

| label                               | Term ID   | L... |   |   |
|-------------------------------------|-----------|------|---|---|
| Entire skin AND subcutaneous tissue | 361714009 | 🌐    | X | 🗨 |

+ Find term  
+ Add term

#### Histopathology

| label | Term ID | L... |  |  |
|-------|---------|------|--|--|
|-------|---------|------|--|--|

+ Find term



# International Classification of Traditional Medicine

---

- Is one of the WHO Family of International Classifications (ICD is another one)
- Will be developed in OWL
- Will use a custom-tailored version of WebProtégé

# Trends in Biomedical Ontologies

---

- A large number of **prominent large** ontologies
- Large **functional repositories** of ontologies
- New ontologies being developed with the Semantic Web principles and technologies at their core
- Legacy internationally developed terminologies adopting Semantic Web principles

# Outline

---

- Biomedical ontologies
- Applications of Semantic Web technologies in biomedicine
- Organizations and consortia promoting Semantic Web technologies in biomedicine
- Challenges and research issues

I've developed an ontology, now what?



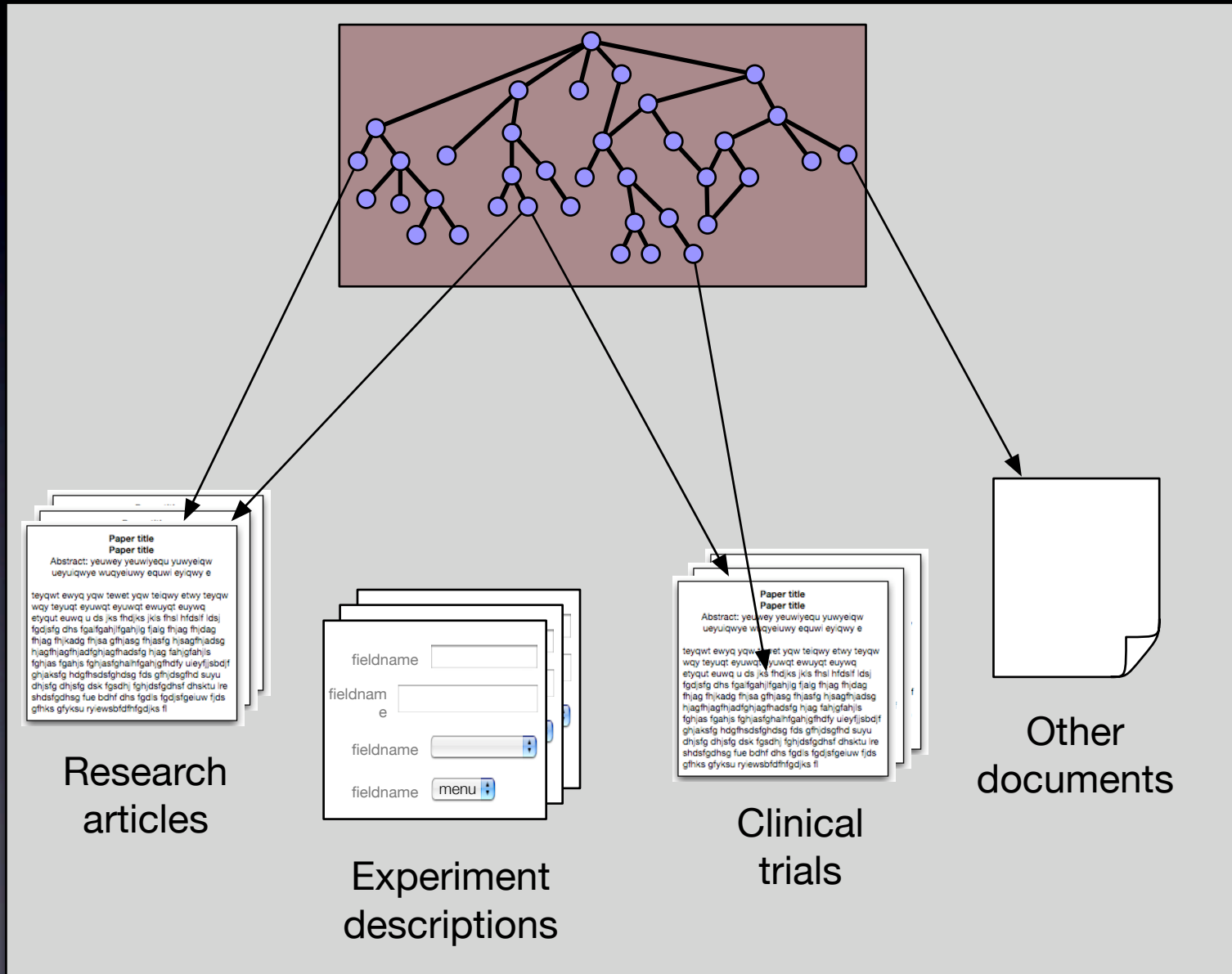
# Applications of Semantic Web Technologies in Biomedicine

---

- Using ontology terms to **name things**: annotation
  - consistent annotation across datasets (Gene Ontology, GO)
  - image annotation (Foundational Model of Anatomy, FMA)
- **Data exchange format**
  - Standard terms for describing experiments (MGED)
- Driving **natural-language processing** (Geneways)
- **Information integration**
  - Homogeneous layer over heterogeneous resources (Transparent Access to Multiple Bioinformatics Information Sources, TAMBIS)



# Annotation



# Manual Annotation: GO Annotations in AmiGO

▼ **Filter tree view** ?

Filter by ontology

Ontology

- All
- biological process
- cellular component
- molecular function

Filter Gene Product Counts

Data source

- All
- ASAP
- AspGD
- CGD

Species

- All
- Arabidopsis thaliana
- Bacillus anthraci...
- Bacillus subtilis

View Options

Tree view  Full  Compact

[Set filters](#)

[Remove all filters](#)

all : all [457986 gene products] [E](#)

- GO:0008150 : biological\_process** [352871 gene products] [E](#)
  - GO:0022610 : biological adhesion** [4776 gene products]
  - GO:0065007 : biological regulation** [64301 gene products]
  - GO:0009758 : carbohydrate utilization** [4 gene products]
  - GO:0015976 : carbon utilization** [172 gene products]
  - induction of necroptosis by extracellular signals ; GO:0060555** [\[show def\]](#) [\[view in tree\]](#)

| Symbol, full name   | Information   | Qualifier   | Reference                      | Assigned by             |
|---|---|---|--------------------------------|-------------------------|
| <input type="checkbox"/> <a href="#">Faslg</a><br>Fas ligand (TNF superfamily, member 6)                  | <a href="#">36 associations</a> <a href="#">BLAST</a> <b>gene</b> from <i>Rattus norvegicus</i> | <a href="#">IEA</a><br>With <a href="#">Ensembl:ENSP00000356694</a> | <a href="#">RGD:1600115</a>    | Ensembl (via RGD)       |
| <input type="checkbox"/> <a href="#">FASLG</a><br>Tumor necrosis factor ligand superfamily member 6       | <a href="#">36 associations</a> <a href="#">BLAST</a> <b>protein</b> from <i>Homo sapiens</i>   | <a href="#">ISO</a><br>With <a href="#">RGD:1606342</a>             | <a href="#">RGD:1624291</a>    | RGD                     |
| <input type="checkbox"/> <a href="#">FASLG</a><br>Tumor necrosis factor ligand superfamily member 6       | <a href="#">36 associations</a> <a href="#">BLAST</a> <b>protein</b> from <i>Homo sapiens</i>   | <a href="#">IDA</a>   | <a href="#">PMID:11101870</a>  | BHF-UCL (via UniProtKB) |
| <input type="checkbox"/> <a href="#">Gzmb</a><br>granzyme B   | <a href="#">9 associations</a> <a href="#">BLAST</a> <b>gene</b> from <i>Rattus norvegicus</i>  | <a href="#">IMP</a>   | <a href="#">RGD:2325193</a>    | RGD                     |
| <input type="checkbox"/> <a href="#">RIPK1</a><br>Receptor (TNFRSF)-interacting serine-threonine kinase 1 | <a href="#">20 associations</a> <a href="#">BLAST</a> <b>protein</b> from <i>Bos taurus</i>     | <a href="#">IEA</a><br>With <a href="#">Ensembl:ENSP00000259808</a> | <a href="#">GO REF:0000019</a> | Ensembl (via UniProtKB) |
| <input type="checkbox"/> <a href="#">Ripk1</a><br>receptor (TNFRSF)-interacting serine-threonine kinase 1 | <a href="#">28 associations</a> <a href="#">BLAST</a> <b>gene</b> from <i>Rattus norvegicus</i> | <a href="#">ISO</a><br>With <a href="#">RGD:1320896</a>             | <a href="#">RGD:1624291</a>    | RGD                     |

# Automatic Annotation: NCBO Annotator

The screenshot displays the NCBO Annotator interface. At the top, the 'Open Biomedical Annotator' window shows configuration options: 'Ontologies' set to 'SNOMEDCT' and 'Semantic Types' set to '7,T033,T200,T026,T029,T023,T038,T017,T047,T048,T191,T019,T121,T195,T020,T050'. The 'Annotate Text' radio button is selected. The input text is 'Melanoma is a disease of the melanocytes affecting the bowel and the eye'. An 'Annotate' button is visible.

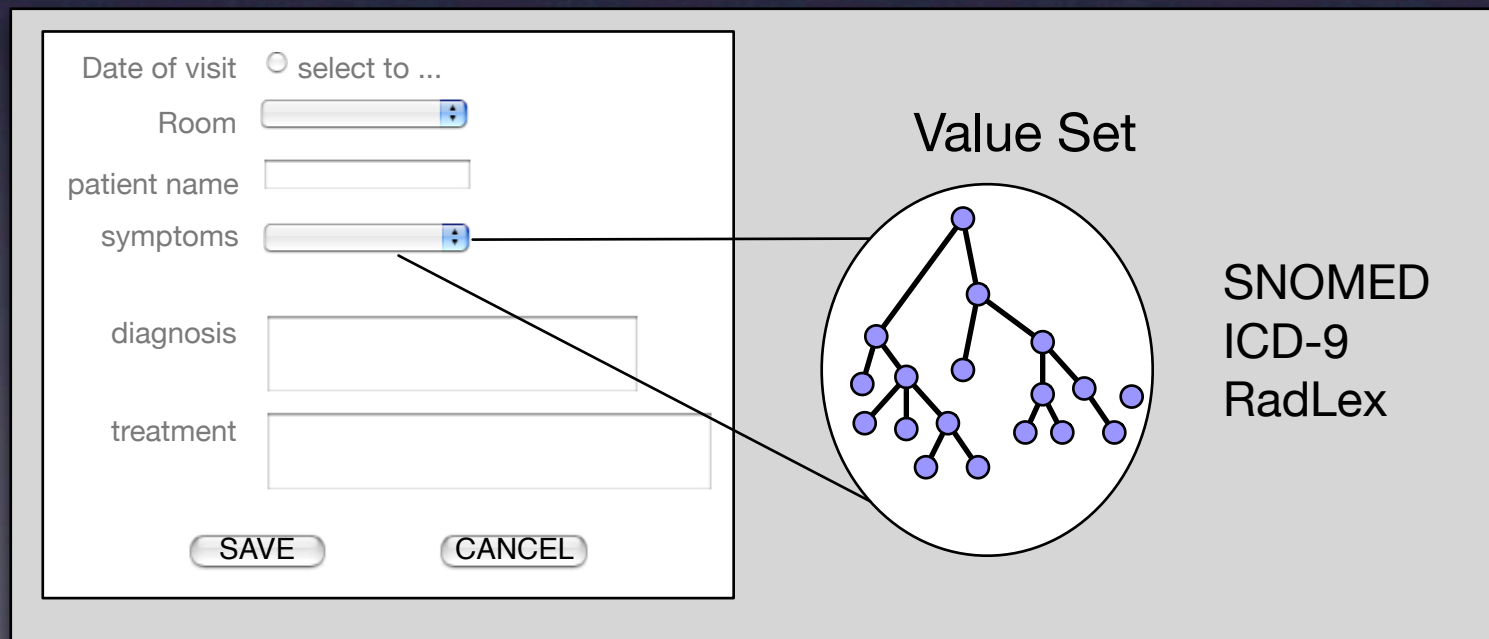
The main interface is divided into three panels:

- Ontologies (1):** Filtered to 'UMLS and BioPortal Ontologies'. A single ontology, 'SNOMED Clinical Terms, 2008\_07\_31', is checked.
- Annotation Tag Cloud (5):** Contains 'Annotation statistics' showing 0 expanded annotations from ISA\_CLOSURE and MAPPING, and 5 direct annotations from MGREP. Below the statistics, the words 'Melanoma', 'Eye', and 'Disease' are displayed in large blue font, with 'Entire eye' and 'Intestines' in smaller blue font.

At the bottom, there are links to learn more about the [NCBO Annotator web service](#) and the [NCBO Biomedical Resources index](#).

# Use of Biomedical Terminologies in Clinical Practice

- Most **Electronic Medical Record (EMR)** systems use controlled terminologies for some of the fields in the record
- They provide a way to integrate information across different EMR systems
  - except that it never really works in practice, because the systems have different configurations (see Research challenges later)
- May enable comparative effectiveness studies, facilitate patient care, etc.





# BioRDF

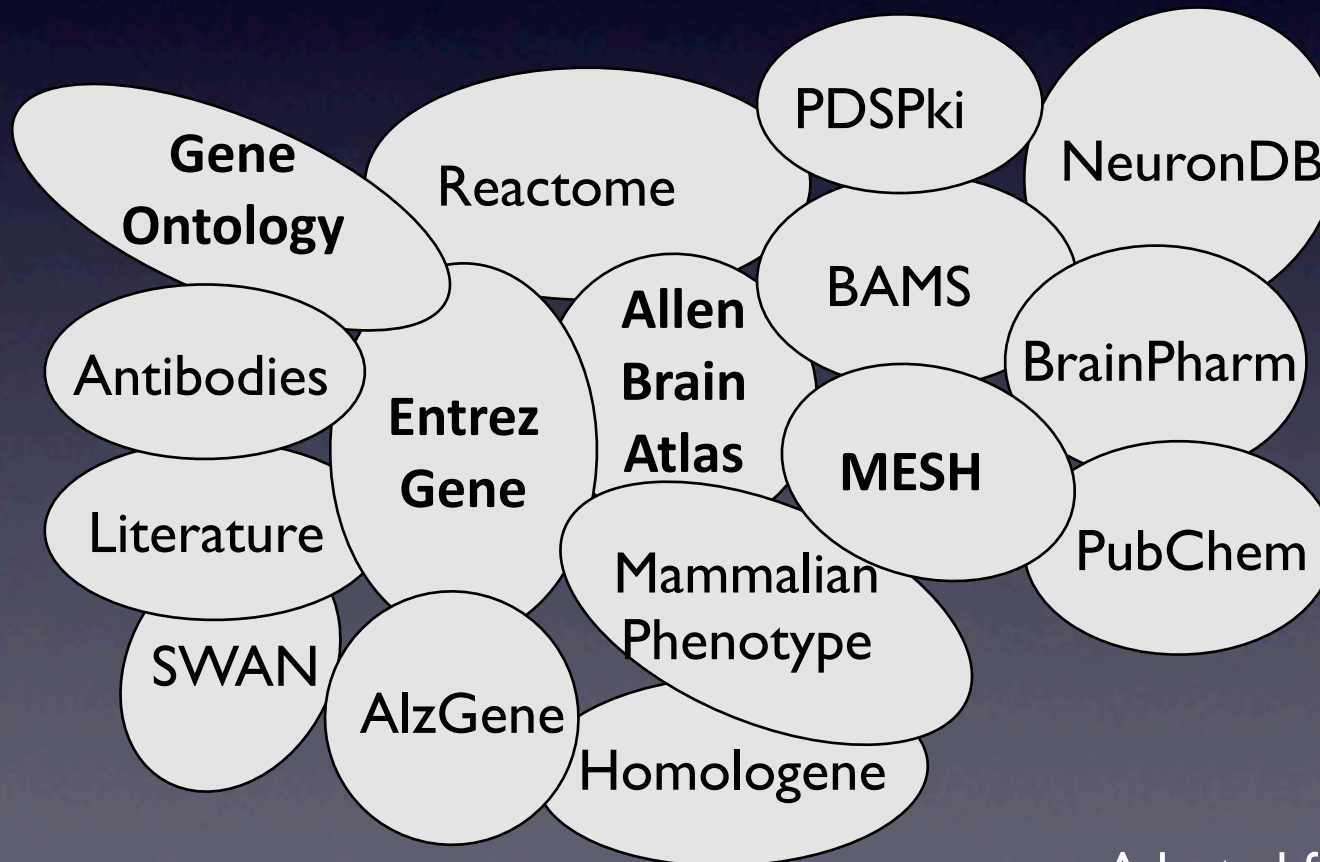
---

- **BioRDF**: An effort by the W3C Health care and Life Sciences Interest Group to represent a large number of biomedical resources in RDF
  - represented several large datasets, mostly relevant to neurological diseases and treatment in RDF
  - used SWObjects to map the data in SQL databases to RDF
  - provided SPARQL access to these resources
  - supported query federation across multiple data points
  - developed demos for specific use cases
    - Receptor explorer: finding receptor-related data

# BioRDF: Integrating Heterogeneous Data

---

- **BioRDF**: An effort by the W3C Health care and Life Sciences Interest Group to represent a large number of biomedical resources in RDF



Adapted from : Susie Stephens

# BioRDF: SPARQL Query

```
prefix go: <http://purl.org/obo/owl/GO#>
prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>
prefix owl: <http://www.w3.org/2002/07/owl#>
prefix mesh: <http://purl.org/commons/record/mesh/>
prefix sc: <http://purl.org/science/owl/sciencecommons/>
prefix ro: <http://www.obofoundry.org/ro/ro.owl#>

select ?genename ?processname
where
{
  graph <http://purl.org/commons/hcls/pubmesh>
  {
    ?paper ?p mesh:D017966 .
    ?article sc:identified_by_pmid ?paper.
    ?gene sc:describes_gene_or_gene_product_mentioned_by ?article.
  }
  graph <http://purl.org/commons/hcls/goa>
  {
    ?protein rdfs:subClassOf ?res.
    ?res owl:onProperty ro:has_function.
    ?res owl:someValuesFrom ?res2.
    ?res2 owl:onProperty ro:realized_as.
    ?res2 owl:someValuesFrom ?process.
  }
  graph <http://purl.org/commons/hcls/20070416/classrelations>
  {
    {?process <http://purl.org/obo/owl/obo#part_of> go:GO_0007166}
    union
    {?process rdfs:subClassOf go:GO_0007166 }
    ?protein rdfs:subClassOf ?parent.
    ?parent owl:equivalentClass ?res3.
    ?res3 owl:hasValue ?gene.
  }
  graph <http://purl.org/commons/hcls/gene>
  { ?gene rdfs:label ?genename }
  graph <http://purl.org/commons/hcls/20070416>
  { ?process rdfs:label ?processname }
}
```

Mesh: Pyramidal Neurons



Pubmed: Journal Articles



Entrez Gene: Genes



GO: Signal Transduction

*Inference required*

Source: Alan Ruttenberg

# BioRDF: Results: Genes, Processes

|                |   |
|----------------|---|
| •DRD1, 1812    | adenylate cyclase activation  |
| •ADRB2, 154    | adenylate cyclase activation  |
| •ADRB2, 154    | arrestin mediated desensitization of G-protein coupled receptor protein signaling pathway |
| •DRD1IP, 50632 | dopamine receptor signaling pathway   |
| •DRD1, 1812    | dopamine receptor; adenylylase activating pathway   |
| •DRD2, 1813    | dopamine receptor; adenylylase inhibiting pathway   |
| •GRM7, 2917    | G-protein coupled receptor protein signaling pathway                                      |
| •GNG3, 2785    | G-protein coupled receptor protein signaling pathway                                      |
| •GNG12, 55970  | G-protein coupled receptor protein signaling pathway                                      |
| •DRD2, 1813    | G-protein coupled receptor protein signaling pathway                                      |
| •ADRB2, 154    | G-protein coupled receptor protein signaling pathway                                      |
| •CALM3, 808    | G-protein coupled receptor protein signaling pathway                                      |
| •HTR2A, 3356   | G-protein coupled receptor protein signaling pathway                                      |
| •DRD1, 1812    | G-protein signaling, coupled to cyclic nucleotide second messenger                        |
| •SSTR5, 6755   | G-protein signaling, coupled to cyclic nucleotide second messenger                        |
| •MTNR1A, 4543  | G-protein signaling, coupled to cyclic nucleotide second messenger                        |
| •CNR2, 1269    | G-protein signaling, coupled to cyclic nucleotide second messenger                        |
| •HTR6, 3362    | G-protein signaling, coupled to cyclic nucleotide second messenger                        |
| •GRIK2, 2898   | glutamate signaling pathway   |
| •GRIN1, 2902   | glutamate signaling pathway   |
| •GRIN2A, 2903  | glutamate signaling pathway   |
| •GRIN2B, 2904  | glutamate signaling pathway   |
| •ADAM10, 102   | integrin-mediated signaling pathway   |
| •GRM7, 2917    | negative regulation of adenylylase activity   |
| •LRP1, 4035    | negative regulation of Wnt receptor signaling pathway                                     |
| •ADAM10, 102   | Notch receptor processing   |
| •ASCL1, 429    | Notch signaling pathway   |
| •HTR2A, 3356   | serotonin receptor signaling pathway  |
| •ADRB2, 154    | transmembrane receptor protein tyrosine kinase activation (dimerization)                  |
| •PTPRG, 5793   | transmembrane receptor protein tyrosine kinase signaling pathway                          |
| •EPHA4, 2043   | transmembrane receptor protein tyrosine kinase signaling pathway                          |
| •NRTN, 4902    | transmembrane receptor protein tyrosine kinase signaling pathway                          |
| •CTNND1, 1500  | Wnt receptor signaling pathway  |

Many of the genes are related to AD through gamma secretase (presenilin) activity

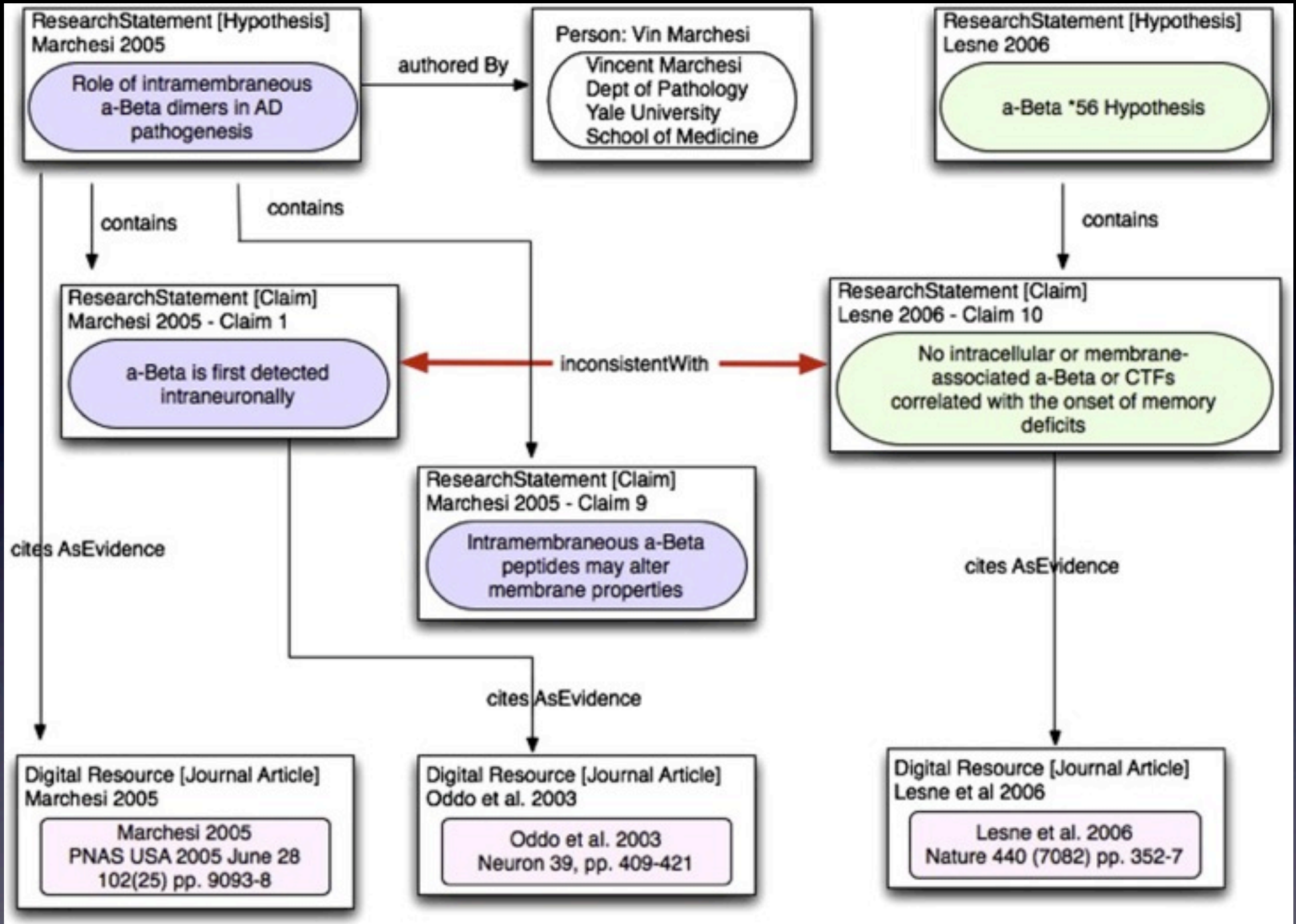
Adapted from: Alan Ruttenberg



# SWAN: Semantic Web Applications in Neuromedicine



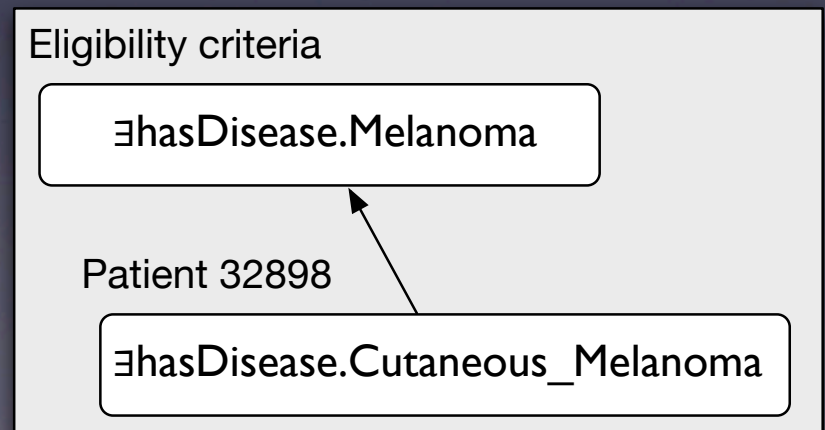
- Using Semantic Web to support **scientific discourse**.
- Specifically applied for Alzheimer's, Parkinson's and other **neurodegenerative diseases**.
- The SWAN ontology supports the application that enables scientists
  - to **formulate scientific hypotheses** and describe them formally
  - to **point to literature** that supports or refutes the hypothesis
  - to **link** different hypotheses
  - to **search** the current hypotheses and their supporting documentation
- Integrates contributions from scientists from different sub-disciplines relevant to Alzheimer's and Parkinson's diseases:
  - psychiatry, neurology, microscopic anatomy, neuronal physiology, biochemistry, genetics, molecular biology, and bioinformatics
- Scientists can **understand** the current state of the art and **ask the right questions** in their experiments



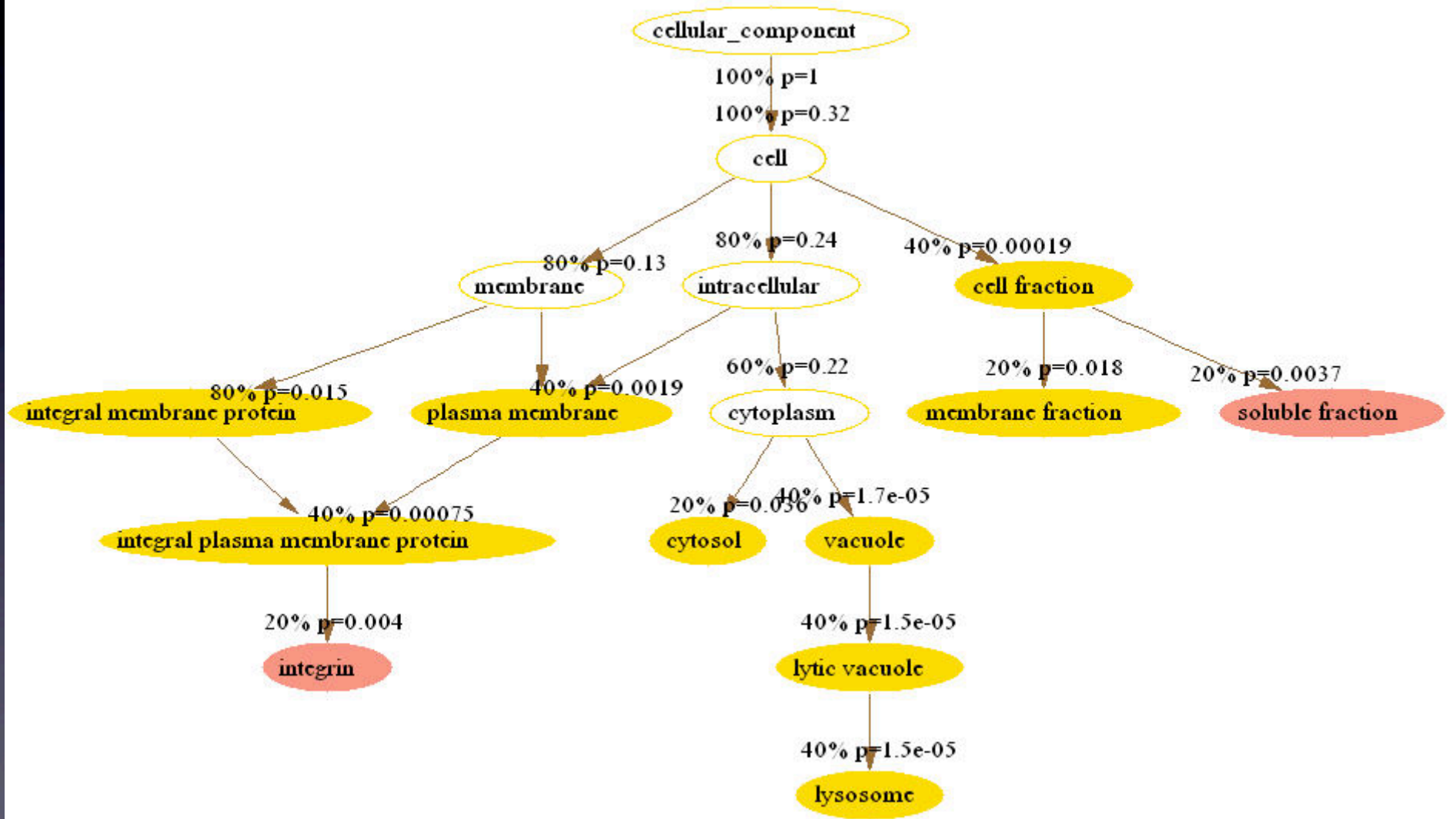
Ciccarese P, Wu E, Kinoshita J, Wong G, Ocana M, Ruttenberg A, Clark T. **The SWAN Biomedical Discourse Ontology.** J Biomed Inform. 2008 Oct;41(5):739-51. Epub 2008 May 4.. PMID: 18583197

# Using OWL Inference: Eligibility Criteria for Clinical Trials

- All clinical trials of new therapies have **eligibility criteria** for participation. For example
  - “a breast cancer patient, who had a mastectomy and has been treated with Taxol for 4 weeks”
- **Matching patients to clinical trials** is a difficult problem: many trials fail because there are not enough patients.
- Medical records describe specific subclasses of drugs, morphology, diagnosis, etc.
  - OWL reasoning can help match such expression describing a patient to an expression defining eligibility criteria



# GO-Enrichment analysis





# Summary: Applications

---

- Annotation of data with ontology terms (manual and automatic)
- Value sets for forms in EMRs and other data
- Query federation across multiple datasets using RDF
- Support for scientific discourse
- Use of inference for eligibility criteria matching
- .... many others



# Outline

---

- Biomedical ontologies
- Applications of Semantic Web technologies in biomedicine
- Organizations and consortia promoting Semantic Web technologies in biomedicine
- Challenges and research issues

# Efforts and Organizations



- W3C Interest Group: Semantic Web Health Care and Life Sciences (HCLS)
  - produces use cases for applications of semantic web technologies in healthcare and life sciences
    - use cases are often specific scenarios in a subdomain (e.g., a researcher trying to perform new experiments to learn about therapies for Alzheimer)
  - BioRDF: “rdfize” a number of diverse resources to showcase the use of Semantic Web technologies
  - Ontology development: Translational Medicine Ontology

# CWA: Concept Web Alliance

---



- Goal: create the Concept Web
  - collaboratively created
  - dynamic
  - graph of concepts and their relationships
- Not really an ontology
- A graph of concepts that can be used and referred to in ontologies

# OBO Foundry



## OBO Foundry ontologies

| Title  | Domain              | Prefix | File                                   | Last changed |
|--|---------------------|--------|--|--------------|
| <a href="#">Biological process</a>                       | biological process  | GO     | <a href="#">gene_ontology_edit.obo</a> | 2010/08/29   |
| <a href="#">Cellular component</a>                       | anatomy             | GO     | <a href="#">gene_ontology_edit.obo</a> | 2010/08/29   |
| <a href="#">Chemical entities of biological interest</a> | biochemistry        | CHEBI  | <a href="#">chebi.obo</a>              | 2010/08/03   |
| <a href="#">Molecular function</a>                       | biological function | GO     | <a href="#">gene_ontology_edit.obo</a> | 2010/08/29   |
| <a href="#">Phenotypic quality</a>                       | phenotype           | PATO   | <a href="#">quality.obo</a>            | 2010/08/19   |
| <a href="#">Protein Ontology (PRO)</a>                   | proteins            | PRO    | <a href="#">pro.obo</a>                | 2010/08/20   |
| <a href="#">Xenopus anatomy and development</a>          | anatomy             | XAO    | <a href="#">xenopus_anatomy.obo</a>    | 2009/12/02   |
| <a href="#">Zebrafish anatomy and development</a>        | anatomy             | ZFA    | <a href="#">zebrafish_anatomy.obo</a>  | 2010/08/06   |

## OBO Foundry candidate ontologies and other ontologies of interest

| Title   | Domain      | Prefix | File                                     | Last changed |
|---|-------------|--------|--|--------------|
| <a href="#">Amphibian gross anatomy</a>       | anatomy     | AAO    | <a href="#">amphibian_anatomy.obo</a>    | 2008/06/19   |
| <a href="#">Amphibian taxonomy</a>            | anatomy     | ATO    | <a href="#">amphibian_taxonomy.obo</a>   |              |
| <a href="#">Ascomycete phenotype ontology</a> | phenotype   | APO    | <a href="#">ascomycete_phenotype.obo</a> | 2010/05/12   |
| <a href="#">Basic Formal Ontology</a>         | upper       | BFO    | <a href="#">1.1</a>                      |              |
| <a href="#">Bilateria anatomy</a>             | anatomy     | BILA   | <a href="#">bilateria_mrca.obo</a>       |              |
| <a href="#">Biological imaging methods</a>    | experiments | FBbi   | <a href="#">image.obo</a>                | 2010/05/05   |
| <a href="#">BRENDA tissue / enzyme source</a> | anatomy     | BTO    | <a href="#">BrendaTissueOBO</a>          |              |



# National Center for Biomedical Ontology

BioPortal [Browse](#) [Search](#) [Projects](#) [Annotate](#) [All Mappings](#) [All Resources Alpha](#) [Sign In](#) [Register](#) [Help/About](#) [Send Feedback](#)

human phenotype ontology

Search all ontologies

 [Search](#)  
[Advanced Search](#)

Find an ontology

 [Explore](#)  
[Browse Ontologies >](#)

Search resources

 [Search](#)  
[Advanced Resource Search](#)

Most Active Ontologies

| Ontology  | Version | Notes | Mappings |
|---|---------|-------|----------|
| <a href="#">Human disease</a>                     | 1.36    | 0     | 17732    |
| <a href="#">Mouse adult gross anatomy</a>         | 1.194   | 0     | 3905     |
| <a href="#">NCI Thesaurus</a>                     | 08.12d  | 9     | 3798     |
| <a href="#">Foundational Model of Anatomy</a>     | 3.0     | 0     | 1997     |
| <a href="#">Zebrafish anatomy and development</a> | 1.21    | 0     | 791      |

Statistics

|                   |         |
|-------------------|---------|
| Ontologies        | 143     |
| Concepts          | 723,806 |
| Resources Indexed | 11      |

Latest Notes

[change namespace from IAO to OBI material entity \(Ontology for Biomedical Investigations\)](#) 06/10/09 whetzel

---

[RE:Add mapping to Unit ontology? unit \(Experimental Factor Ontology\)](#) 06/01/09 jamesmalone

---

[Add mapping to Unit ontology? unit \(Experimental Factor Ontology\)](#) 04/12/09 whetzel

---

[Release notes for version 2009-04-02 entity \(NanoParticle Ontology\)](#) 04/02/09 sobolevnm

---

[Missing preferred term field OBI\\_0000577 \(Ontology for Biomedical Investigations\)](#) 04/01/09 whetzel

Latest Mappings

[cellular\\_component \(Biological process\) => Cell\\_component \(Foundational Model of Anatomy\)](#) 04/27/09 matthiassamwald

---

[Cell\\_component \(Foundational Model of Anatomy\) => cellular\\_component \(Biological process\)](#) 04/27/09 matthiassamwald

---

[Cell \(Foundational Model of Anatomy\) => Cell \(NCI Thesaurus\)](#) 04/16/09 lechatpito

---

[Cell \(NCI Thesaurus\) => Cell \(Foundational Model of Anatomy\)](#) 04/16/09 lechatpito

---

[tbio:Cell \(Basic Vertebrate Anatomy\) => Cell \(NCI Thesaurus\)](#) 04/16/09 lechatpito

# Conferences and Workshops

---

- If you want to follow the research on Semantic Web technologies for Biomedicine, look at the proceedings of:
  - ISWC
  - International Conference on Biomedical Ontology (ICBO)
  - Bio-ontologies SIG at the International Conference on Intelligent Systems for Molecular Biology (ISMB)
  - Annual Fall Symposium of the American Medical Informatics Association (AMIA)
  - AMIA Translational Bioinformatics Summit
  - Pacific Symposium on Biocomputing (PSB)

# Outline

---

- Biomedical ontologies
- Applications of Semantic Web technologies in biomedicine
- Organizations and consortia promoting Semantic Web technologies in biomedicine
- Challenges and research issues

# Issues and Trends

---

- Ontologies are fairly large and often (fairly) expressive
- Considerable amount of ontology reuse
- Linked data in bioinformatics
- Manual curation of knowledge
  - that feeds back into the ontologies
- Collaborative development of ontologies
  - multiple editors
  - distributed
  - well-defined processes and protocols
  - different workflows



# Some challenges, research issues

---

- **Common ontology language**
  - The OBO format is quite popular in parts of biomedical community
  - The OWL API provides an OBO-to-OWL converter
- **Finding ontologies that are good or useful**
  - ontology evaluation
- **Ontology mappings**
  - lots of overlap, in part because there are lots of legacy ontologies
  - current algorithms are not scalable to the size of these ontologies
- **Collaborative ontology development**
- **Shared value sets to facilitate data integration**