

# Navigating the Translational Researcher Through A Complex of Animal and Biological Resources

NIH, March 6, 2006

## Ontological Resources for the Translational Researcher



*Olivier Bodenreider*

Lister Hill National Center  
for Biomedical Communications  
Bethesda, Maryland - USA

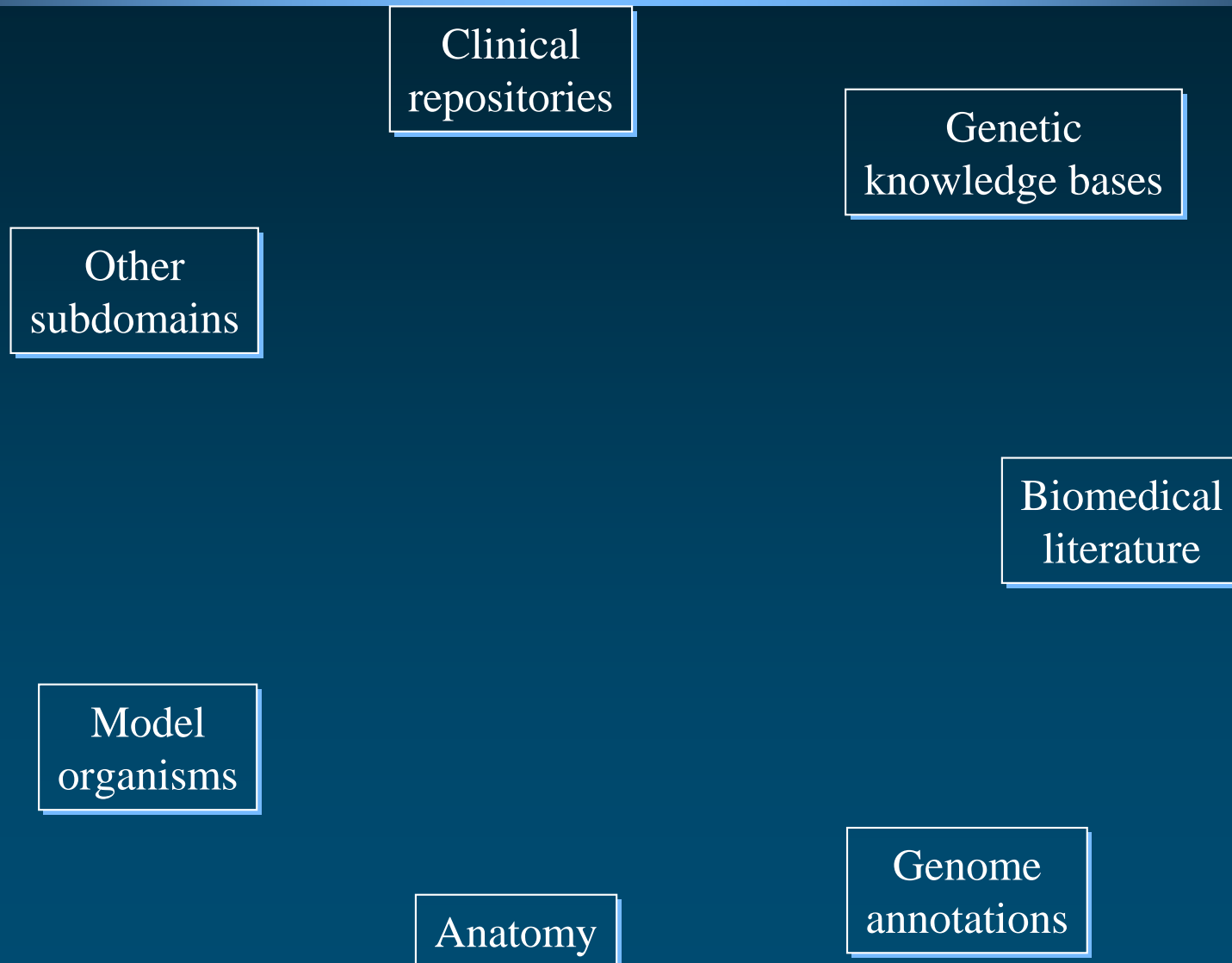
# Overview

- ◆ Information integration
  - Through terminology integration
- ◆ Animals in ontological resources
- ◆ Other (ontological) resources
- ◆ Some issues



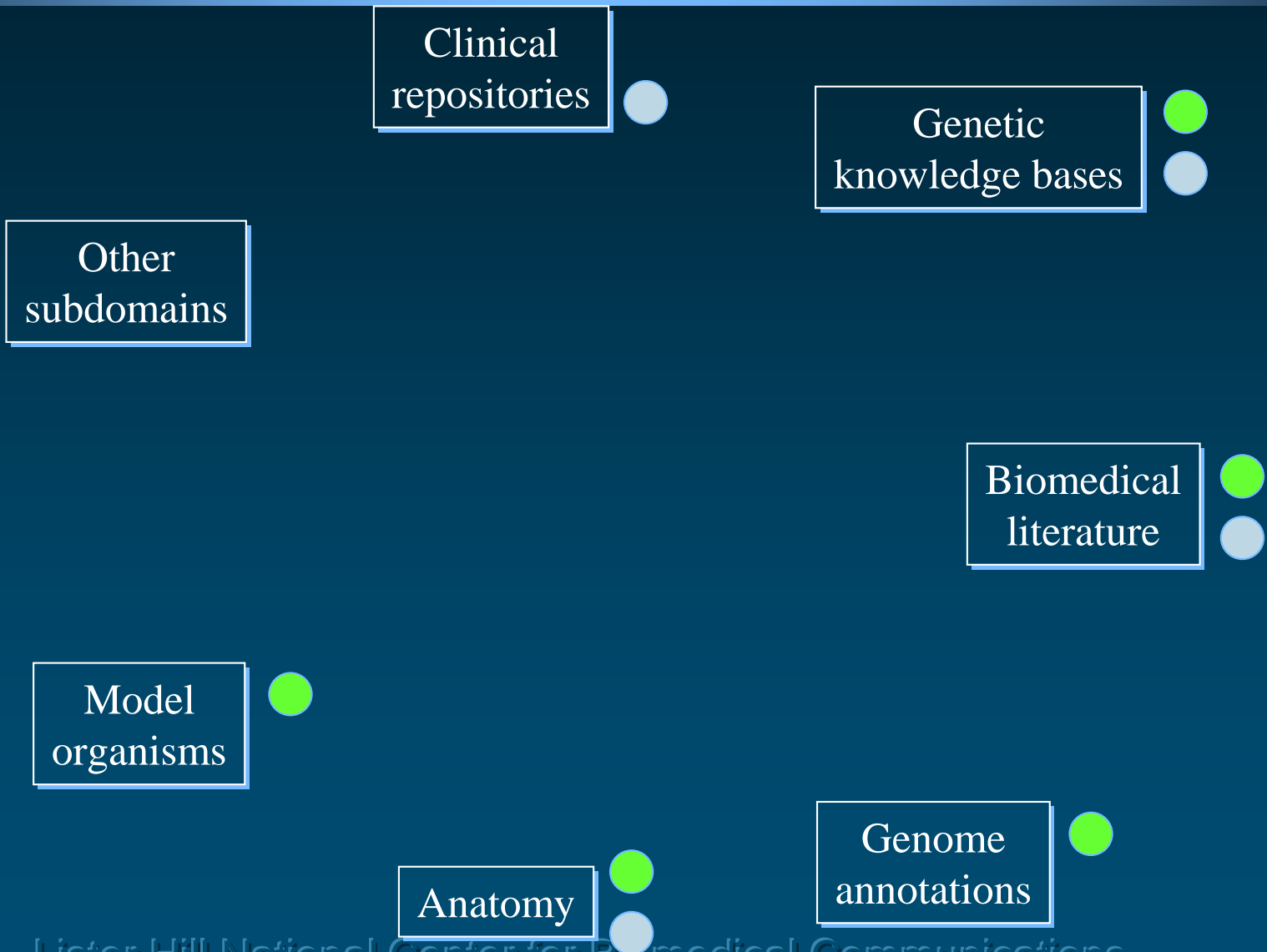
# Information integration

# Subdomains of interest

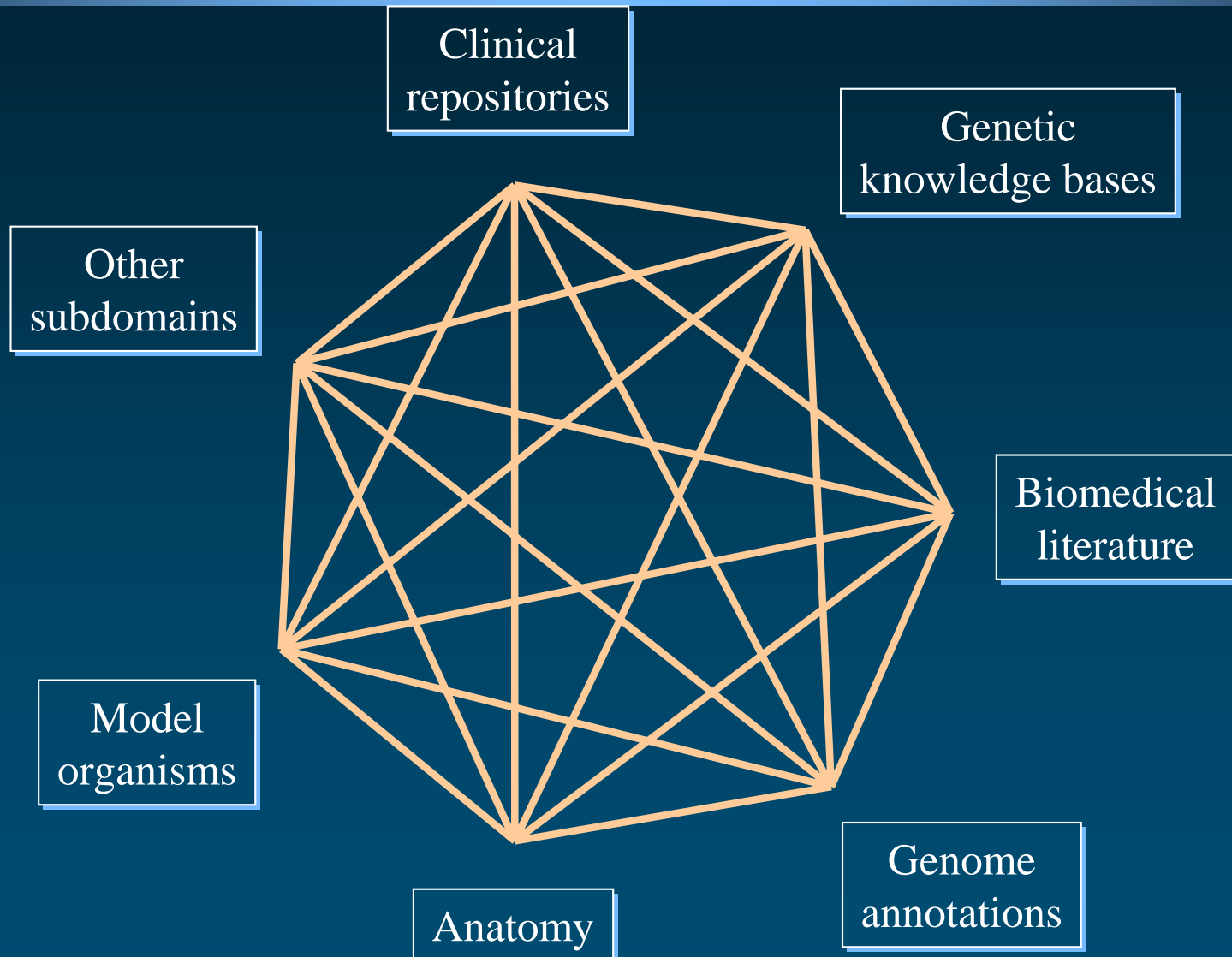


# Subdomains of interest

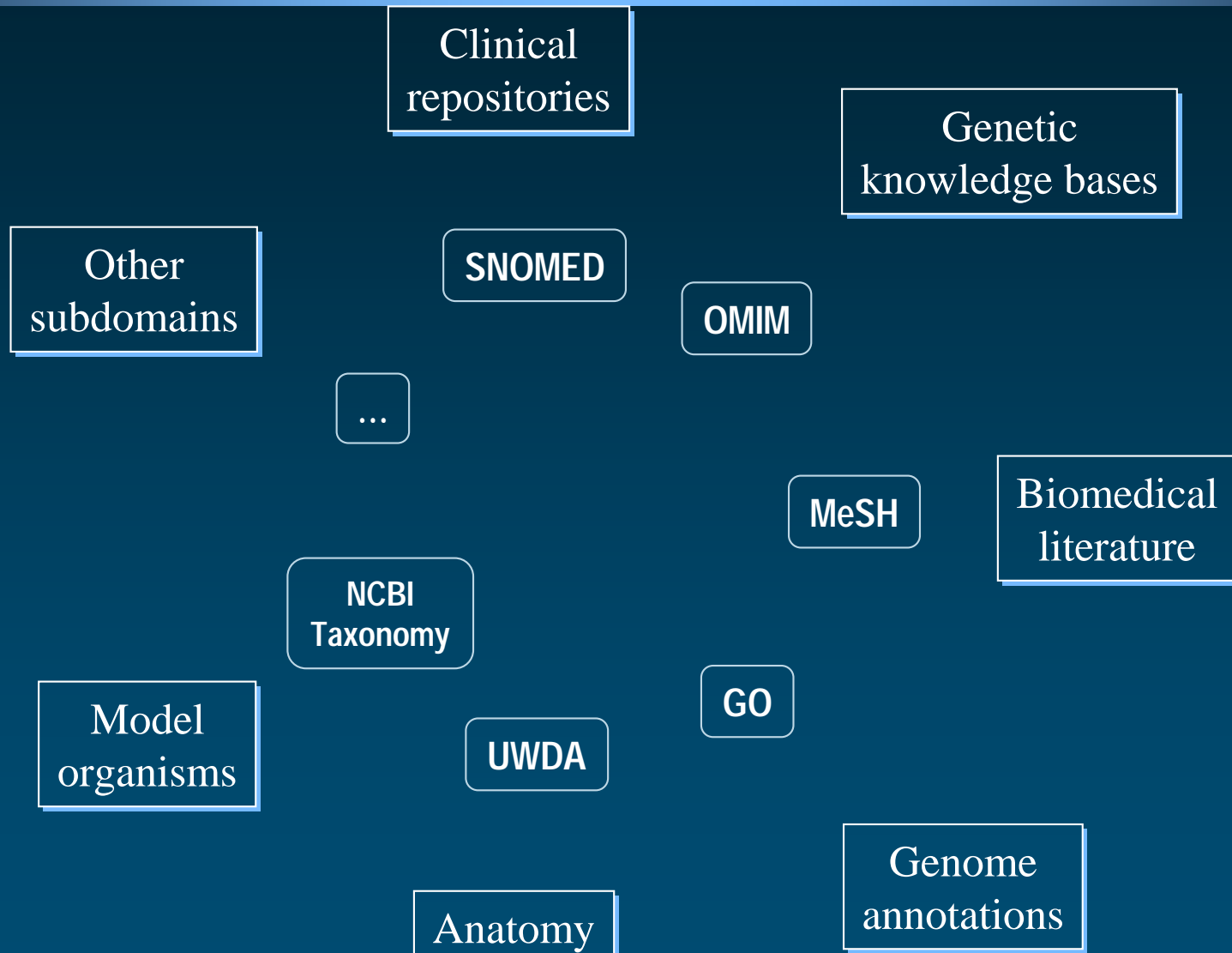
- bench
- bedside



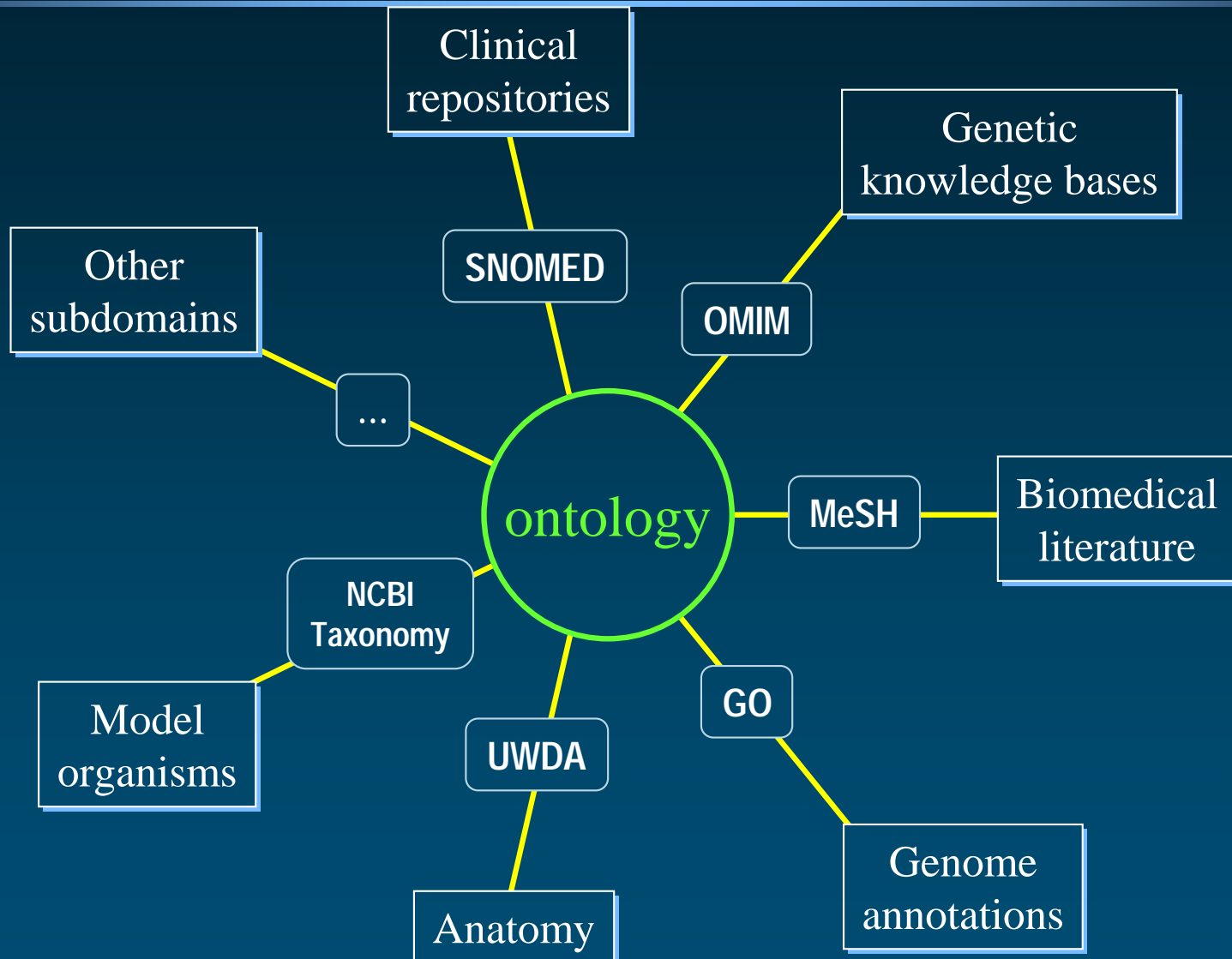
# Integrating subdomains



# Integrating subdomains

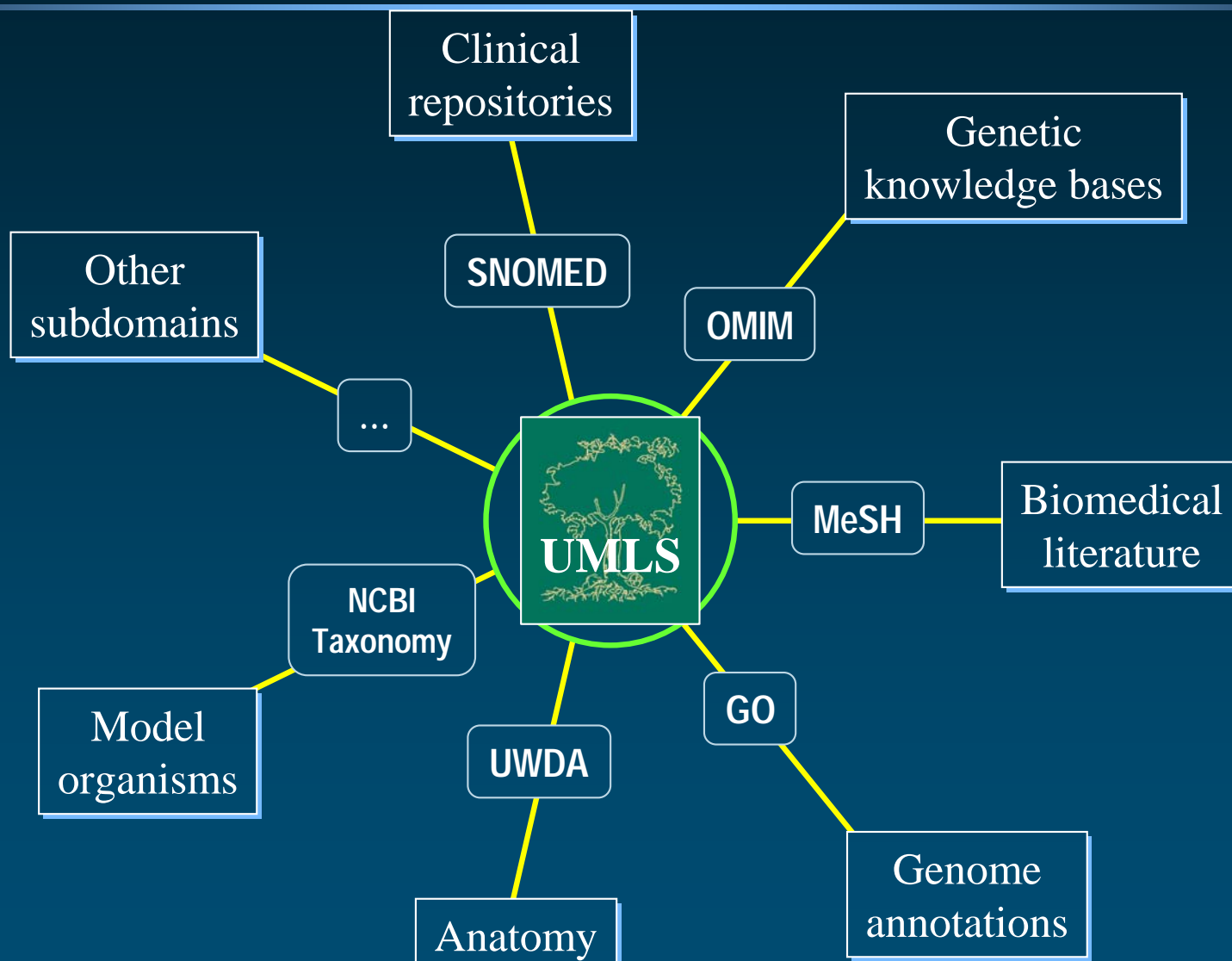


# Integrating subdomains





# Integrating terminologies



# Terminology vs. ontology

## ◆ Terminological resources

- Collections of terms (e.g., controlled vocabularies)
- Useful for indexing and annotation
- MeSH, GO

## ◆ Ontological resources

- Collections of
  - kinds of entities (substances, qualities, processes)
  - relations among them
- Useful for **reasoning**
- UMLS Semantic Network, SNOMED CT



# Animals in ontological resources



## Mus musculus

*Taxonomy ID:* 10090

*Genbank common name:* **house mouse**

*Rank:* species

*Genetic code:* [Translation table 1 \(Standard\)](#)

*Mitochondrial genetic code:* [Translation table 2 \(Vertebrate Mitochondrial\)](#)

*Other names:*

common name: **mouse**

includes: **transgenic mice**

includes: **nude mice**

includes: **LK3 transgenic mice**

includes: **Mus sp. 129SV**

misnomer: **Mus muscaris**

*Lineage( full )*

[cellular organisms](#); [Eukaryota](#); [Fungi/Metazoa group](#); [Metazoa](#); [Eumetazoa](#); [Bilateria](#); [Coelomata](#); [Deuterostomia](#); [Chordata](#); [Craniata](#); [Vertebrata](#); [Gnathostomata](#); [Teleostomi](#); [Euteleostomi](#); [Sarcopterygii](#); [Tetrapoda](#); [Amniota](#); [Mammalia](#); [Theria](#); [Eutheria](#); [Euarchontoglires](#); [Glires](#); [Rodentia](#); [Sciurognathi](#); [Muroidea](#); [Muridae](#); [Murinae](#); [Mus](#)

Entrez records		
Database name	Subtree links	Direct links
<a href="#">Nucleotide</a>	<a href="#">8,144,439</a>	<a href="#">7,804,794</a>
<a href="#">Protein</a>	<a href="#">194,404</a>	<a href="#">193,541</a>
<a href="#">Structure</a>	<a href="#">1,636</a>	<a href="#">1,636</a>
<a href="#">Genome Sequences</a>	<a href="#">45</a>	<a href="#">43</a>
Popset	<a href="#">7,805</a>	<a href="#">7,800</a>
<a href="#">SNP</a>	<a href="#">1,863,252</a>	<a href="#">1,863,252</a>
<a href="#">3D Domains</a>	<a href="#">7,485</a>	<a href="#">7,485</a>
<a href="#">Domains</a>	<a href="#">12</a>	<a href="#">12</a>
<a href="#">GEO Datasets</a>	<a href="#">1,555</a>	<a href="#">1,555</a>
<a href="#">GEO Expressions</a>	<a href="#">5,602,009</a>	<a href="#">5,602,009</a>
<a href="#">UniGene</a>	<a href="#">43,104</a>	<a href="#">43,104</a>
<a href="#">UniSTS</a>	<a href="#">52,676</a>	<a href="#">52,676</a>
<a href="#">PubMed Central</a>	<a href="#">1,800</a>	<a href="#">1,647</a>
<a href="#">HomoloGene</a>	<a href="#">20,520</a>	<a href="#">20,520</a>
<a href="#">Taxonomy</a>	<a href="#">10</a>	<a href="#">1</a>

# Animals in the biomedical literature

## ◆ Check tags in MEDLINE citations

- General: **Animals**, Comparative Study, **Humans**, In Vitro, Pregnancy
- Specific animals

Cattle	Hamsters
Chick Embryo	Mice
Dogs	Rabbits
Guinea Pigs	Rats

- Others: Age groups, Gender, Chronologic tags



# Example from MEDLINE

Michael McClain R, Wolz E, Davidovich A, Bausch J.

Genetic toxicity studies with genistein.

Food Chem Toxicol. 2006 Jan;44(1):42-55. Epub 2005 Sep 28.

PMID: 16198038 [PubMed - indexed for MEDLINE]



## MeSH Terms:

- ◆ [Administration, Oral](#)
- ◆ [Animals](#)
- ◆ [Anticarcinogenic Agents/classification](#)
- ◆ [Anticarcinogenic Agents/toxicity\\*](#)
- ◆ [Bone Marrow Cells/drug effects](#)
- ◆ [Bone Marrow Cells/pathology](#)
- ◆ [Cell Line, Tumor](#)
- ◆ [Female](#)
- ◆ [Genistein/classification](#)
- ◆ [Genistein/toxicity\\*](#)
- ◆ [Leukemia L5178/drug therapy](#)
- ◆ [Leukemia L5178/enzymology](#)
- ◆ [Leukemia L5178/genetics](#)

- ◆ [Male](#)
- ◆ [Mice](#)
- ◆ [Mice, Inbred Strains](#)
- ◆ [Micronuclei, Chromosome-Defective/chemically induced](#)
- ◆ [Micronucleus Tests](#)
- ◆ [Mutagenicity Tests\\*](#)
- ◆ [Mutagens/classification](#)
- ◆ [Mutagens/toxicity\\*](#)
- ◆ [Mutation](#)
- ◆ [Rats](#)
- ◆ [Rats, Wistar](#)
- ◆ [Salmonella typhimurium/drug effects](#)
- ◆ [Salmonella typhimurium/genetics](#)
- ◆ [Salmonella typhimurium/metabolism](#)
- ◆ [Thymidine Kinase/deficiency](#)
- ◆ [Thymidine Kinase/genetics](#)



# Animals in MeSH

## ▶ [Animals \[B01\]](#)

[Animal Population Groups \[B01.050\]](#) +  
[Chordata \[B01.150\]](#) +  
[Invertebrates \[B01.500\]](#) +

## [Animals \[B01\]](#)

[Animal Population Groups \[B01.050\]](#)

[Animals, Laboratory \[B01.050.199\]](#)

[Animals, Inbred Strains \[B01.050.199.520\]](#)

[Animals, Congenic \[B01.050.199.520.5\]](#)

▶ [Mice, Inbred Strains \[B01.050.199.520.5.1\]](#)

[Mice, Inbred A \[B01.050.199.520.5.1.1\]](#)

[Mice, Inbred AKR \[B01.050.199.520.5.1.2\]](#)

## [Animals \[B01\]](#)

[Chordata \[B01.150\]](#)

[Vertebrates \[B01.150.900\]](#)

[Mammals \[B01.150.900.649\]](#)

[Rodentia \[B01.150.900.649.865\]](#)

[Muridae \[B01.150.900.649.865.635\]](#)

[Murinae \[B01.150.900.649.865.635.505\]](#)

▶ [Mice \[B01.150.900.649.865.635.505.1\]](#)

[Mice, Congenic \[B01.150.900.649.865.635.505.1.5\]](#)

[Mice, Inbred Strains \[B01.150.900.649.865.635.505.1.1\]](#)

[Mice, Mutant Strains \[B01.150.900.649.865.635.505.1.2\]](#)

[Mice, Transgenic \[B01.150.900.649.865.635.505.1.3\]](#)

[Rats \[B01.150.900.649.865.635.505.2\]](#)

# Mice in MeSH

[Mice, Congenic \[B01.150.900.649.865.635.505.500.150\]](#)

[Mice, Inbred Strains \[B01.150.900.649.865.635.505.500.400\]](#)

[Mice, Inbred A \[B01.150.900.649.865.635.505.500.400.300\]](#)

[Mice, Inbred AKR \[B01.150.900.649.865.635.505.500.400.318\]](#)

[Mice, Inbred BALB C \[B01.150.900.649.865.635.505.500.400.338\]](#)

[Mice, Inbred C3H \[B01.150.900.649.865.635.505.500.400.388\]](#)

...

[Mice, Mutant Strains \[B01.150.900.649.865.635.505.500.550\]](#)

[Mice, Biozzi \[B01.150.900.649.865.635.505.500.550.100\]](#)

[Mice, Inbred mdx \[B01.150.900.649.865.635.505.500.550.265\]](#)

[Mice, Jimpy \[B01.150.900.649.865.635.505.500.550.430\]](#)

[Mice, Knockout \[B01.150.900.649.865.635.505.500.550.455\]](#)

[Mice, Neurologic Mutants \[B01.150.900.649.865.635.505.500.550.480\]](#)

[Mice, Nude \[B01.150.900.649.865.635.505.500.550.500\]](#)

[Mice, Obese \[B01.150.900.649.865.635.505.500.550.530\]](#)

[Mice, Quaking \[B01.150.900.649.865.635.505.500.550.560\]](#)

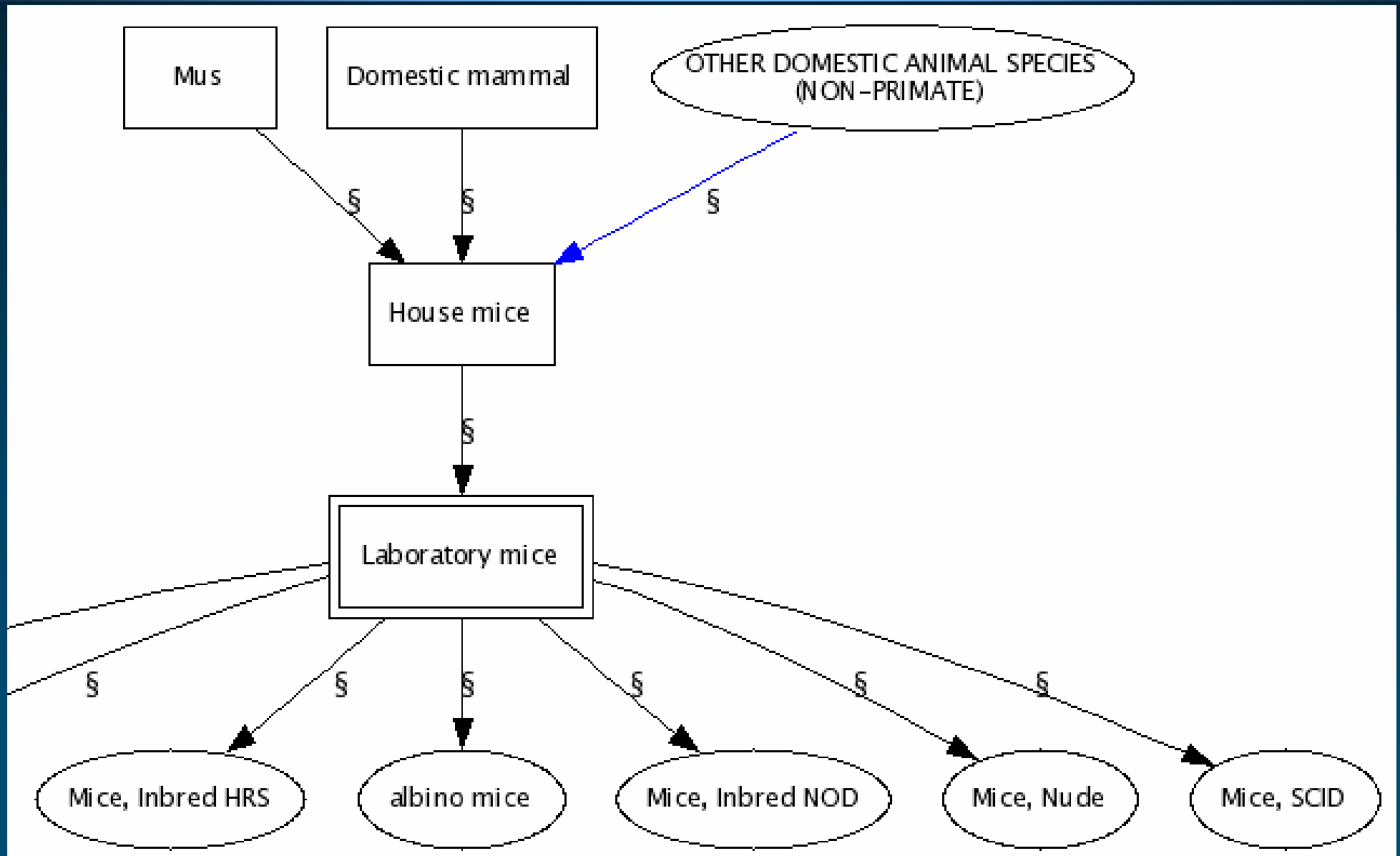
[Mice, SCID \[B01.150.900.649.865.635.505.500.550.780\]](#)

[Mice, Transgenic \[B01.150.900.649.865.635.505.500.800\] +](#)

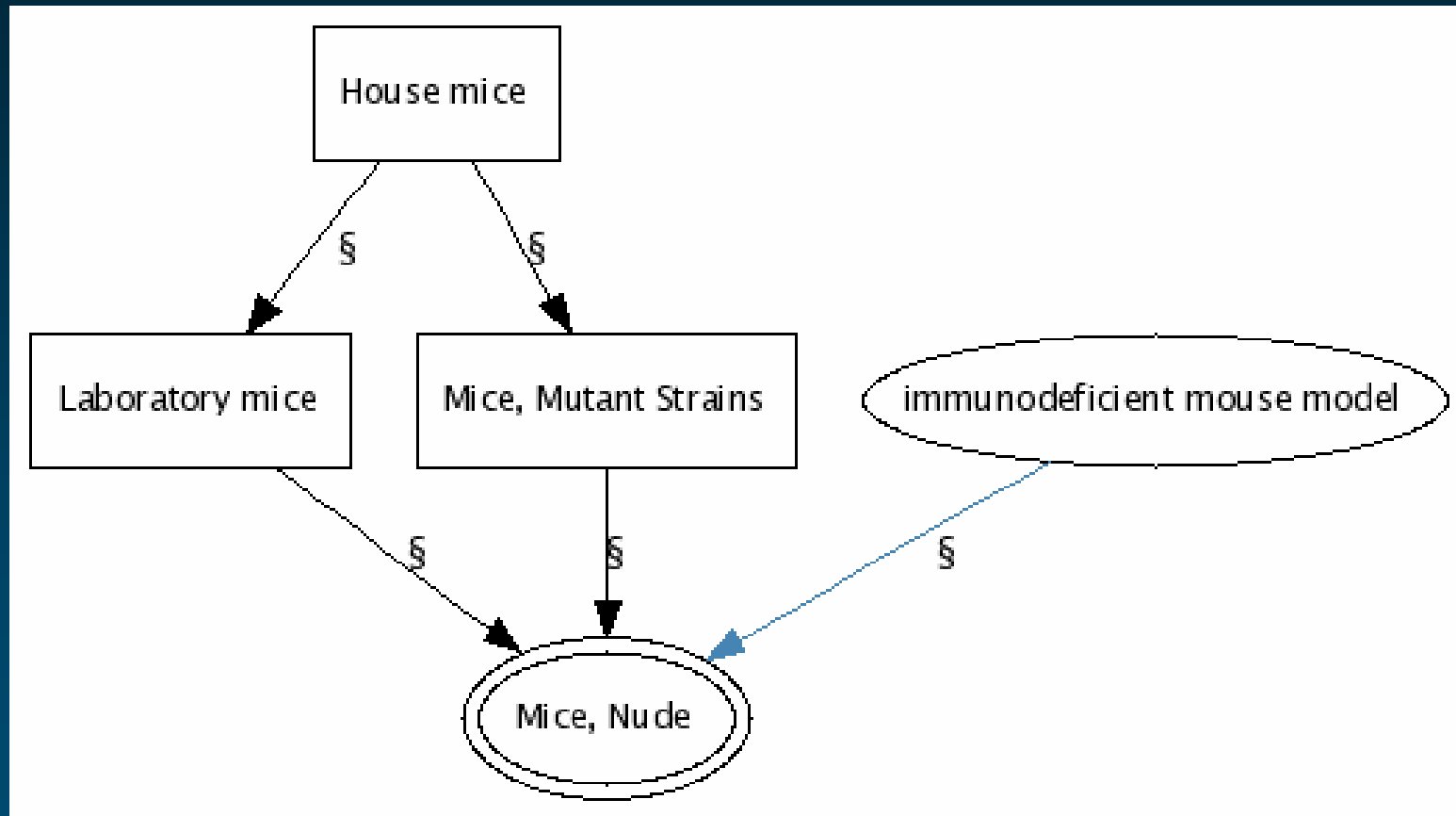




# Mice in UMLS



# Mice in UMLS



Other (ontological) resources

# Open Biological Ontologies



- ◆ Extended family of the Gene Ontology (GO)
- ◆ Collaborative development
  - <http://obo.sourceforge.net/>
- ◆ National Center for Biomedical Ontology
  - <http://bioontology.org/>



- [-] **anatomy**
  - [-] **cell type**
  - [-] **gross anatomy**
    - [-] **animal gross anatomy**
      - [-] C. elegans gross anatomy
      - [-] Drosophila gross anatomy
      - [-] eVoc (Expressed Sequence Annotation for Humans)
    - [-] **fish anatomy**
      - [-] Medaka fish anatomy and development
      - [-] Zebrafish anatomy and development
    - [-] **human developmental anatomy**
      - [-] Human developmental anatomy, timed
      - [-] Human developmental anatomy, abstract
    - [-] Mosquito gross anatomy
    - [-] **mouse anatomy**
      - [-] Mouse adult anatomy
      - [-] Mouse anatomy and development
  - [-] **microbial anatomy**
    - [-] Fungal anatomy
    - [-] Dictyostelium anatomy
  - [-] **plant anatomy**
    - [-] Arabidopsis anatomy
    - [-] Cereal anatomy
    - [-] Maize anatomy
    - [-] Plant anatomy

[-] **development**

[-] **animal development**

- ➔ C. elegans development
- ➔ Drosophila development
- ➔ eVOC (Expressed Sequence Annotation for Humans)
- ➔ medaka fish anatomy and development
- ➔ Mouse anatomy and development
- ➔ Plasmodium life cycle
- ➔ Zebrafish anatomy and development

[-] **plant development**

- ➔ plant growth and developmental stage
- ➔ Arabidopsis development
- ➔ cereal plant development

[-] **phenotype**

- ➔ cereal plant trait
- ➔ context of phenotype
- ➔ eVOC (Expressed Sequence Annotation for Humans)
- ➔ human diseases
- ➔ mammalian phenotype
- ➔ medaka fish anatomy and development
- ➔ mouse pathology
- ➔ plant environmental conditions
- ➔ phenotype attributes and values

Some issues

# Ontology vs. formalism

- ◆ Ontology languages
  - OWL
  - Protégé
- ◆ Markup languages  
(format/syntax for exchanging data)
  - CellML
  - MAGE-ML





# Ontology and granularity

- ◆ The information represented in most ontologies may not be fine-grained enough for some biological applications
  - Strain
  - ...
- ◆ Ontologies: represent **classes** of entities  
vs.  
Biological experiments: refer to **instances**

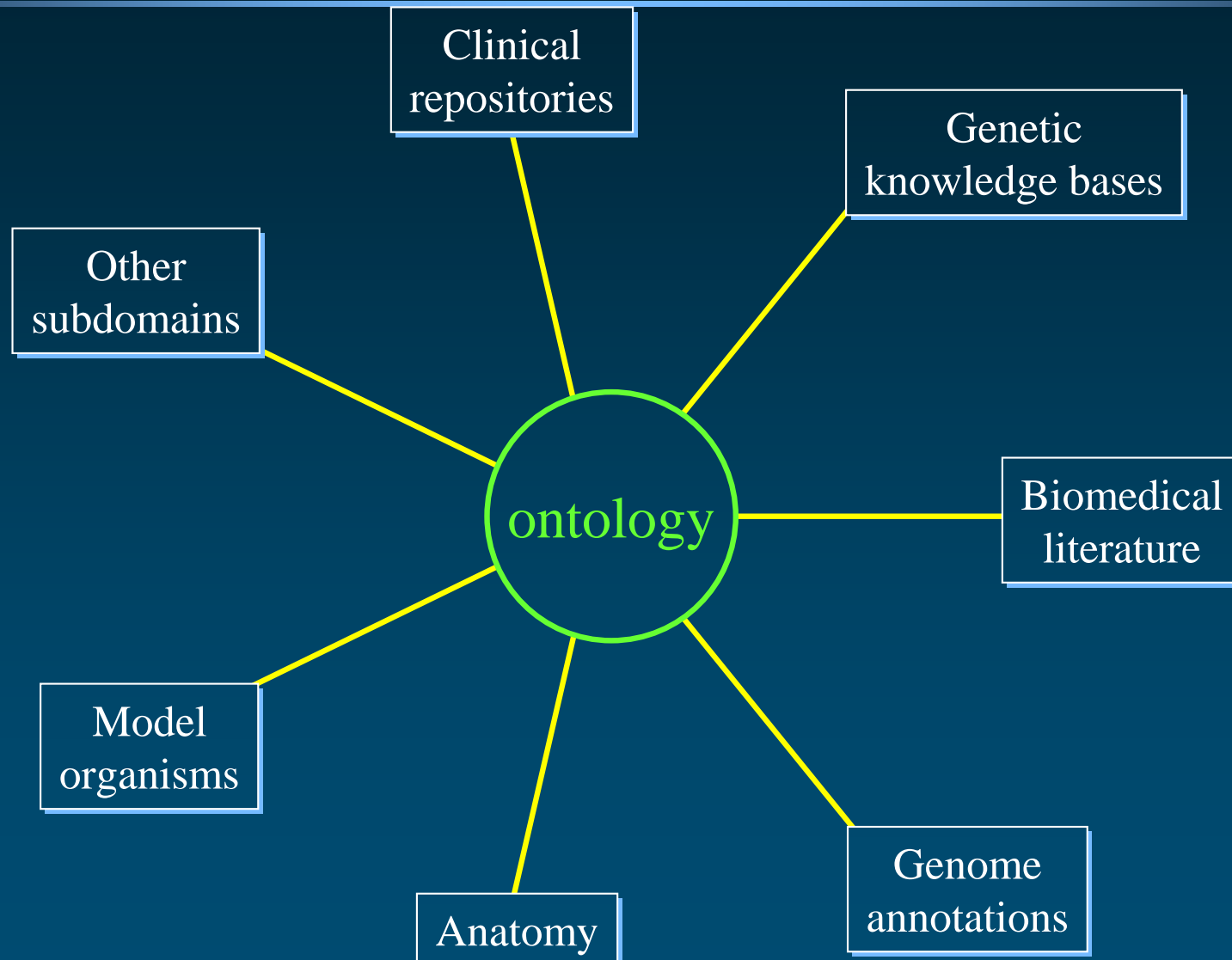
# Phenotype ontologies

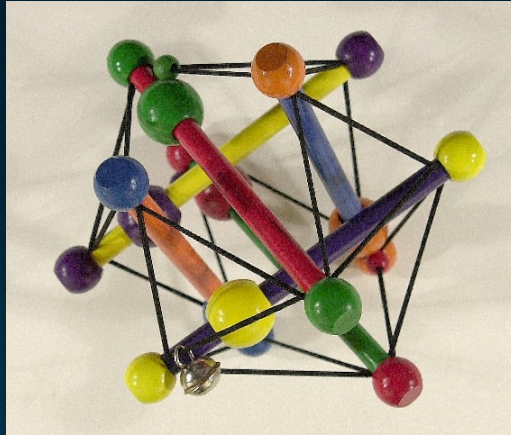
- ◆ Emerging ontologies
- ◆ Fine-grained
  - Attributes
    - Relative Age
    - Carbohydrate Concentration
  - Values
    - Round
    - Tactile Hyperresponsive
- ◆ Multiple organisms

Examples from PATO



# Information integration through ontology





# Medical Ontology Research

Contact: [olivier@nlm.nih.gov](mailto:olivier@nlm.nih.gov)

Web: [mor.nlm.nih.gov](http://mor.nlm.nih.gov)



*Olivier Bodenreider*

Lister Hill National Center  
for Biomedical Communications  
Bethesda, Maryland - USA