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

Clinical repositories

Genetic knowledge bases

Other subdomains

Biomedical literature

Model organisms

Genome annotations

Anatomy

Subdomains of interest

- bench
- bedside

Clinical repositories

Genetic knowledge bases

Other subdomains

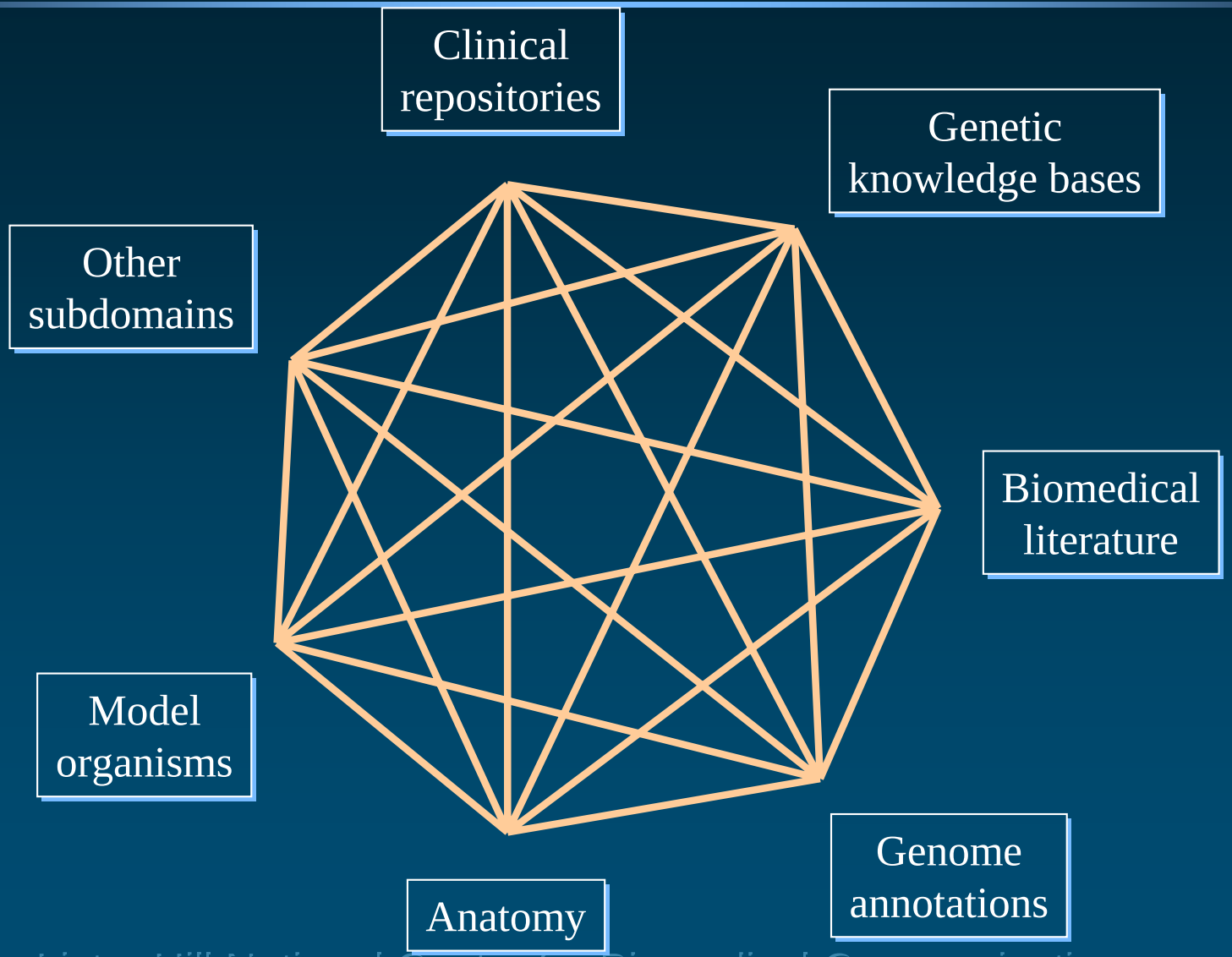
Biomedical literature

Model organisms

Genome annotations

Anatomy

Integrating subdomains



Integrating subdomains

Clinical repositories

Genetic knowledge bases

Other subdomains

SNOMED

OMIM

...

MeSH

Biomedical literature

NCBI Taxonomy

GO

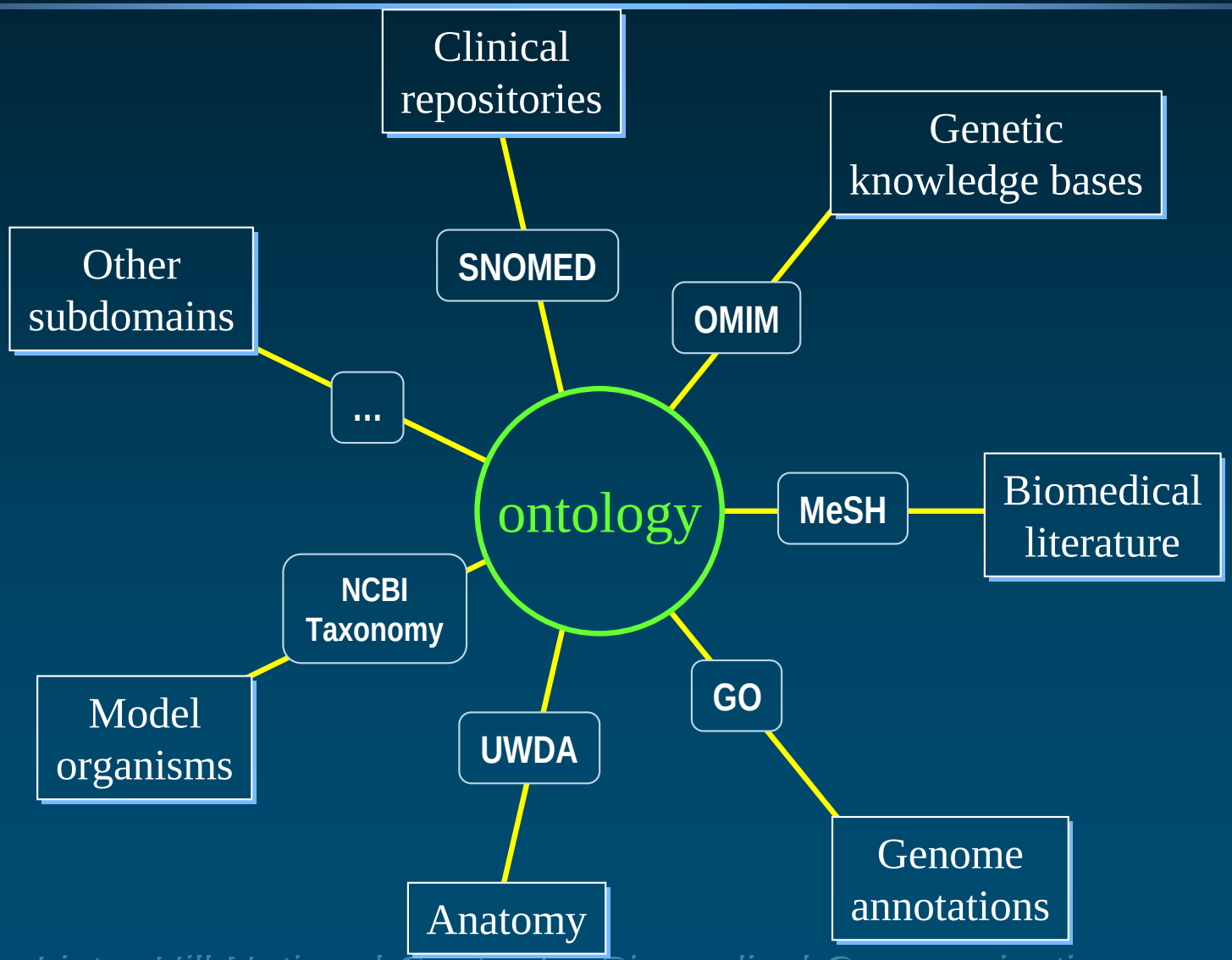
Model organisms

UWDA

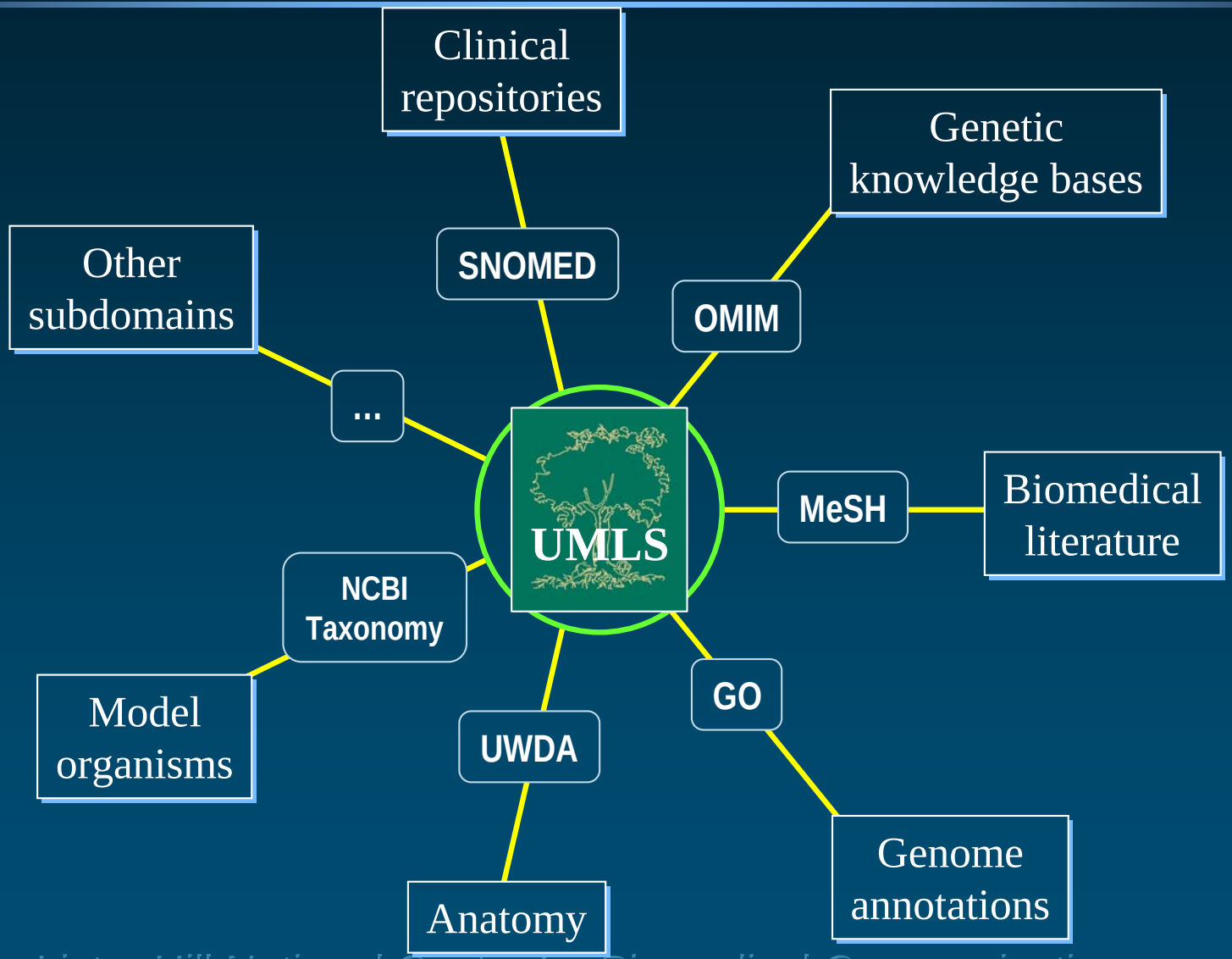
Genome annotations

Anatomy

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

Ontological gradient

Animals in ontological resources

NCBI

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**

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

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

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 [B01.050.19
        - Mice, Inbred Strains [B01.050.
        - Mice, Inbred A [B01.
        - Mice, Inbred AKR [B

- 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 [B01.150.900.649.865.63
              - Mice, Congenic [B01.
              - Mice, Inbred Strains [
              - Mice, Mutant Strains [
              - Mice, Transgenic [B0
              - Rats [B01.150.900.649.865.635

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]

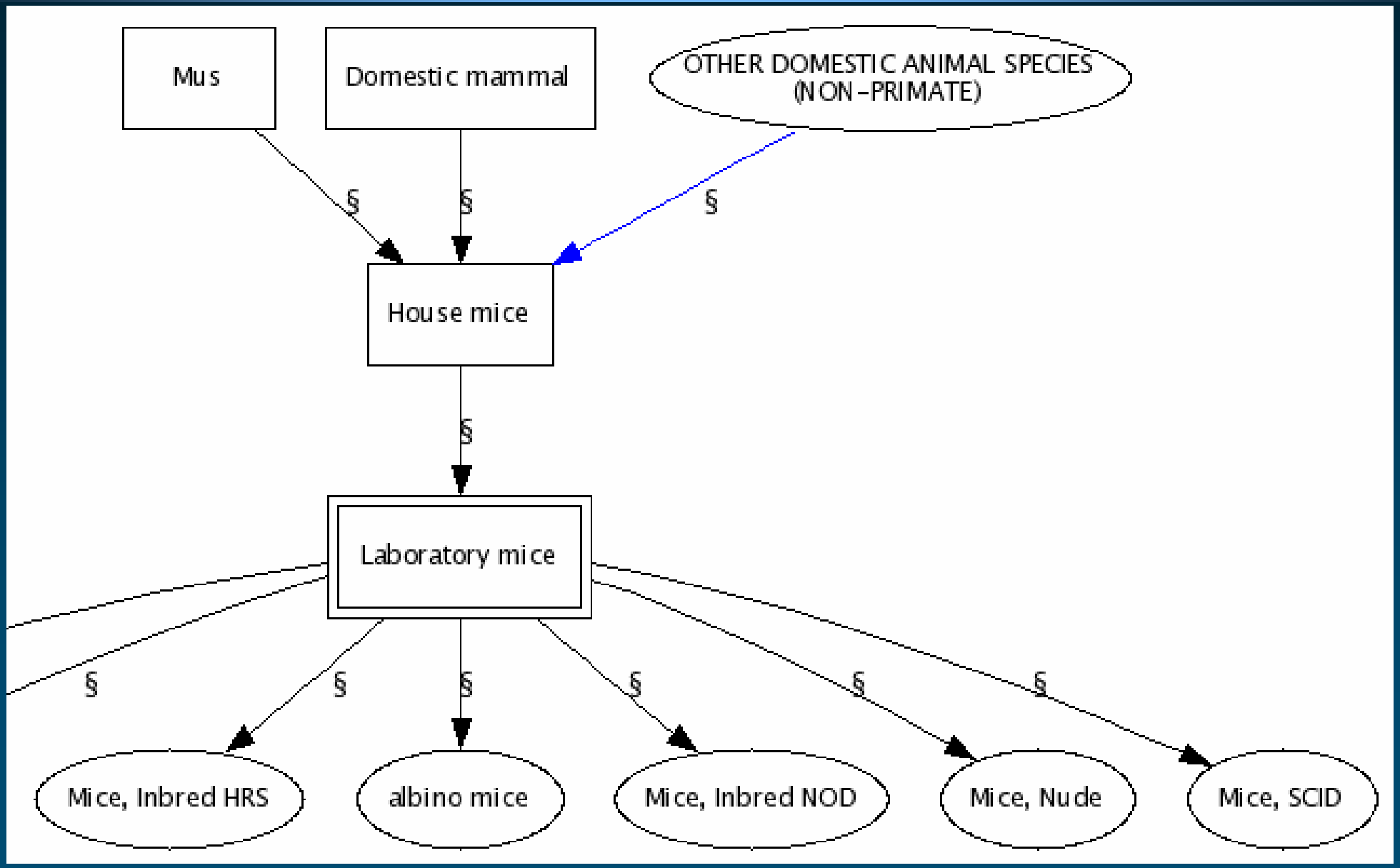
...

Mice, Mutant Strains [B01.150.900.649.865.635.505.500.550]

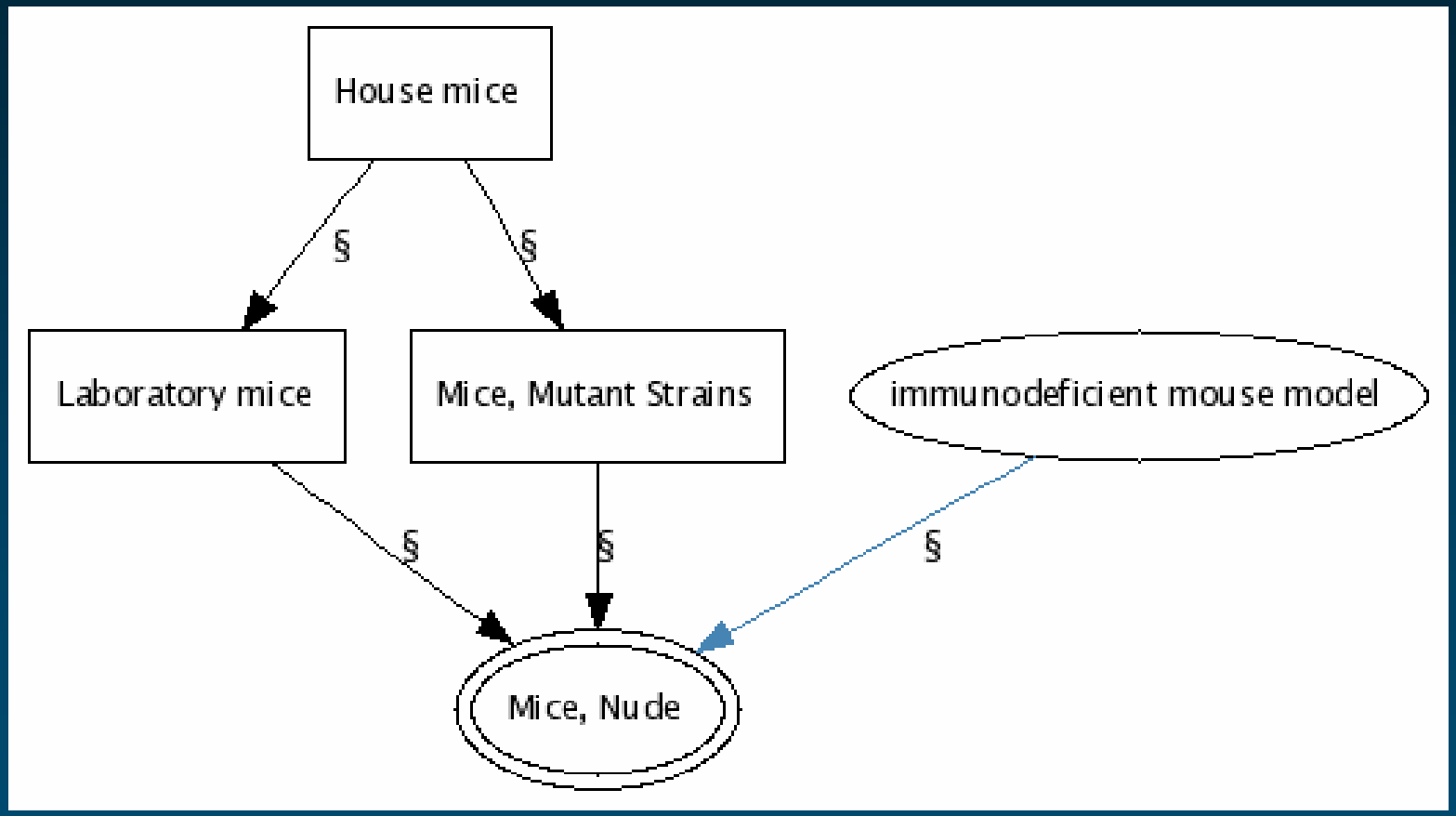
- Mice, Biozzi [B01.150.900.649.865.635.505.500.550.100]
- Mice, Inbred mdx [B01.150.900.649.865.635.505.500.550.265]
- Mice, Jimpy [B01.150.900.649.865.635.505.500.550.430]
- Mice, Knockout [B01.150.900.649.865.635.505.500.550.455]
- Mice, Neurologic Mutants [B01.150.900.649.865.635.505.500.550.480]
- 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


Mus
Domestic mammal
OTHER DOMESTIC ANIMAL SPECIES (NON-PRIMATE)
§
§
§
House mice
§
Laboratory mice
§
§
§
§
§
§
Mice, Inbred HRS
albino mice
Mice, Inbred NOD
Mice, Nude
Mice, SCID

Mice in UMLS


House mice
§
§
Laboratory mice
Mice, Mutant Strains
immunodeficient mouse model
§
§
§
Mice, Nude

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/

OBO
open biomedical ontologies

- 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

OBO
open biomedical ontologies

- 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 enviromental 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

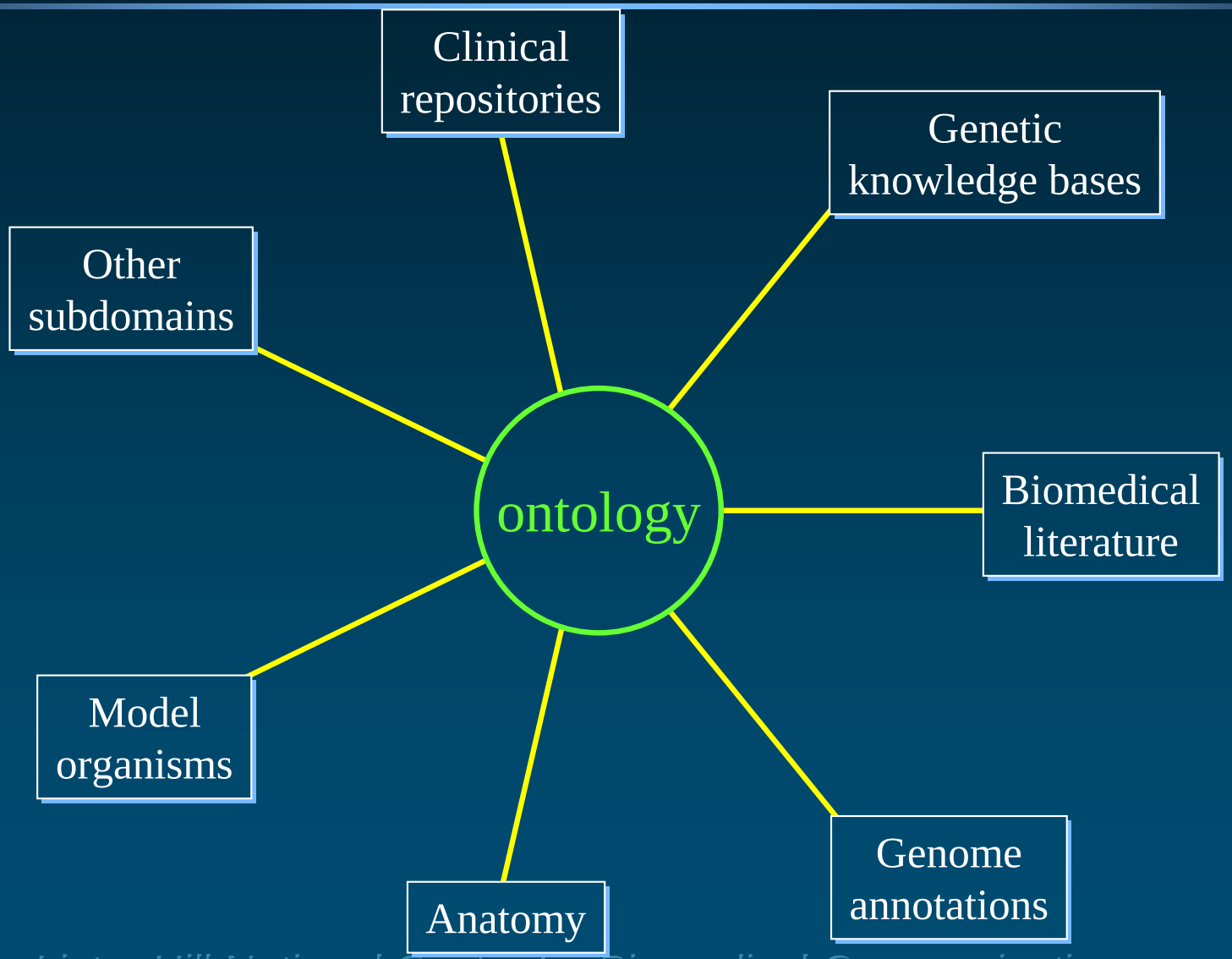
Phenotype ontologies

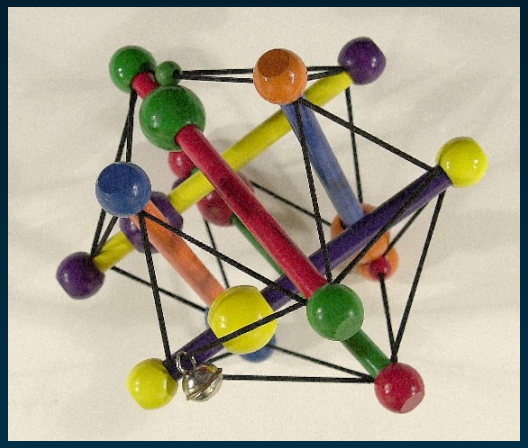
- Emerging ontologies
- Fine-grained
  - Attributes
    - Relative Age
    - Carbohydrate Concentration
  - Values
    - Round
    - Tactile Hyperresponsive

Examples from PATO

- Multiple organisms

Information integration through ontology





Medical Ontology Research

Contact: olivier@nlm.nih.gov

Web: mor.nlm.nih.gov

Olivier Bodenreider

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