



NIAID Ontology Workshop

Focus on Immunology

Gaithersburg, March 21, 2006

UMLS and Semantic Web

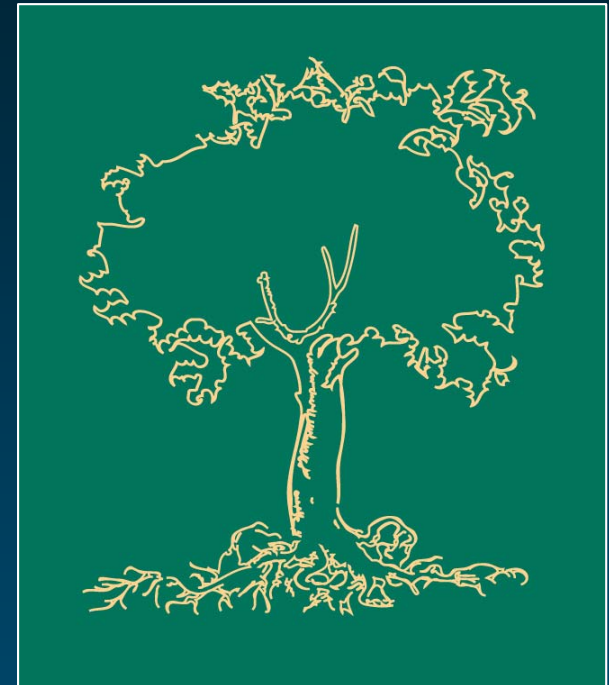


Olivier Bodenreider

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

What does UMLS stand for?

- ◆ Unified
- ◆ Medical
- ◆ Language
- ◆ System



UMLS[®]
Unified Medical Language System[®]
UMLS Metathesaurus[®]



Motivation

- ◆ Started in 1986
- ◆ National Library of Medicine
- ◆ “Long-term R&D project”

«[...] the UMLS project is an effort to overcome two significant barriers to effective retrieval of machine-readable information.

- The first is **the variety of ways the same concepts are expressed** in different machine-readable sources and by different people.
- The second is the **distribution** of useful information among many disparate databases and systems.»

Unified Medical Language System



◆ SPECIALIST Lexicon

- 200,000 lexical items
- Part of speech and variant information

◆ Metathesaurus

- 5M names from over 100 terminologies
- 1M concepts
- 16M relations

◆ Semantic Network

- 135 high-level categories
- 7000 relations among them

Lexical
resources

Terminological
resources

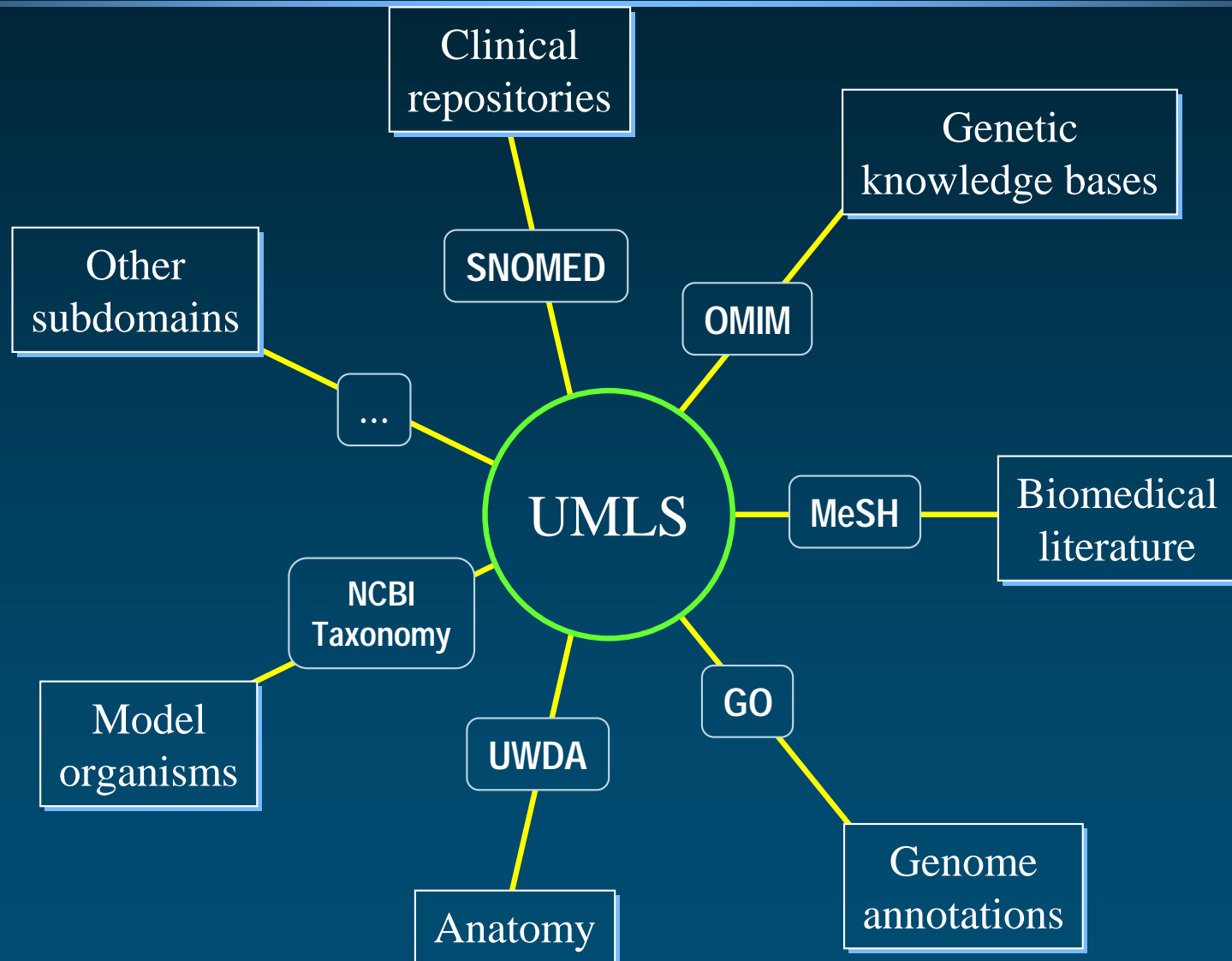
Ontological
resources

Source Vocabularies

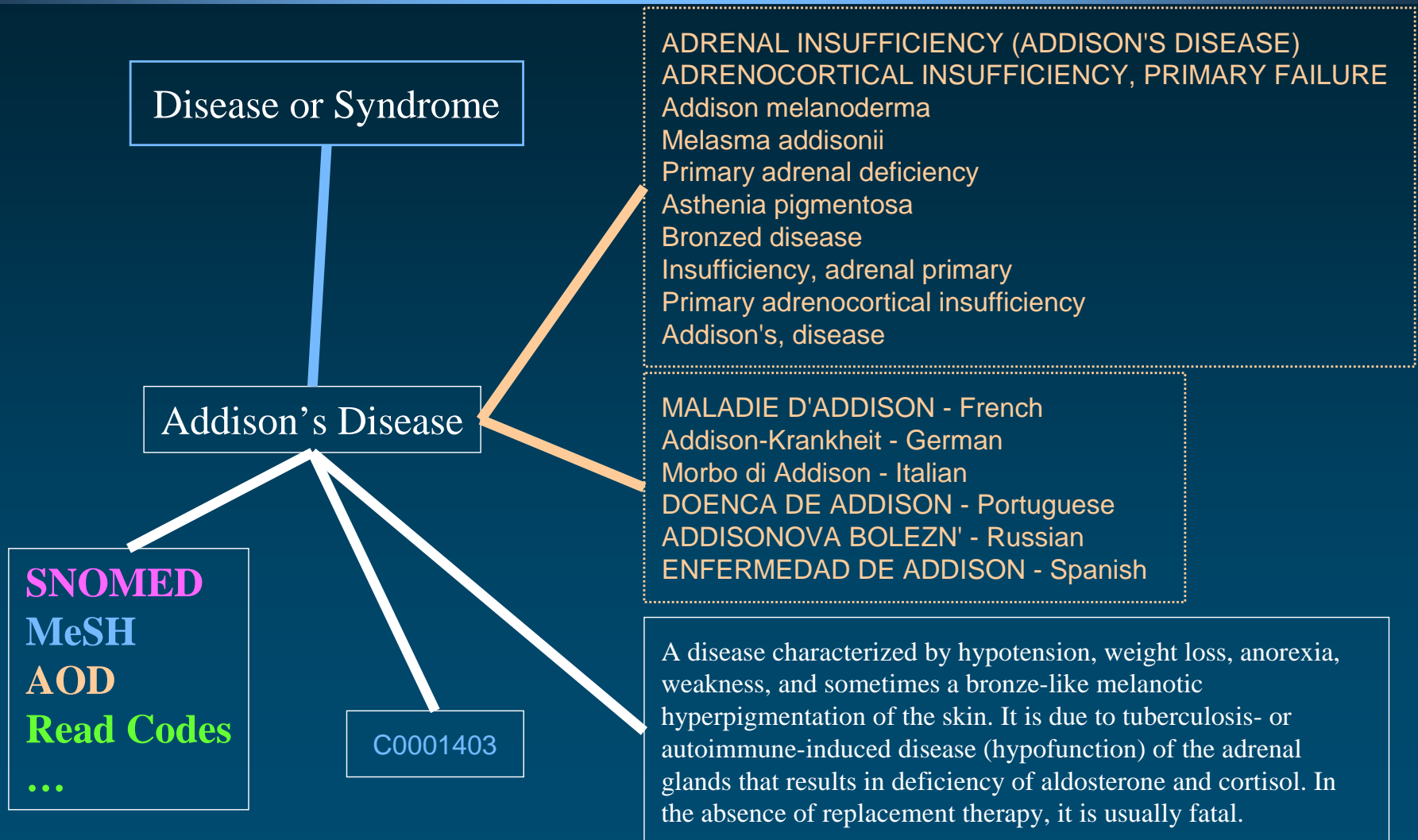
(2006AA)

- ◆ 140 source vocabularies
 - 17 languages
- ◆ Broad coverage of biomedicine
 - 5M names
 - 1.3M concepts
 - 16M relations
- ◆ Common presentation

Integrating subdomains



Addison's Disease: Concept



Metathesaurus Concepts (2005AB)

- ◆ Concept (~ 1.2 M) CUI
 - Set of synonymous concept names
- ◆ Term (~ 4.2 M) LUI
 - Set of normalized names
- ◆ String (~ 4.8 M) SUI
 - Distinct concept name
- ◆ Atom (~ 5.6 M) AUI
 - Concept name in a given source

A0000001 headache (source 1)
A0000002 headache (source 2)
S0000001

A0000003 Headache (source 1)
A0000004 Headache (source 2)
S0000002

L0000001

A0000005 Cephalgia (source 1)
S0000003

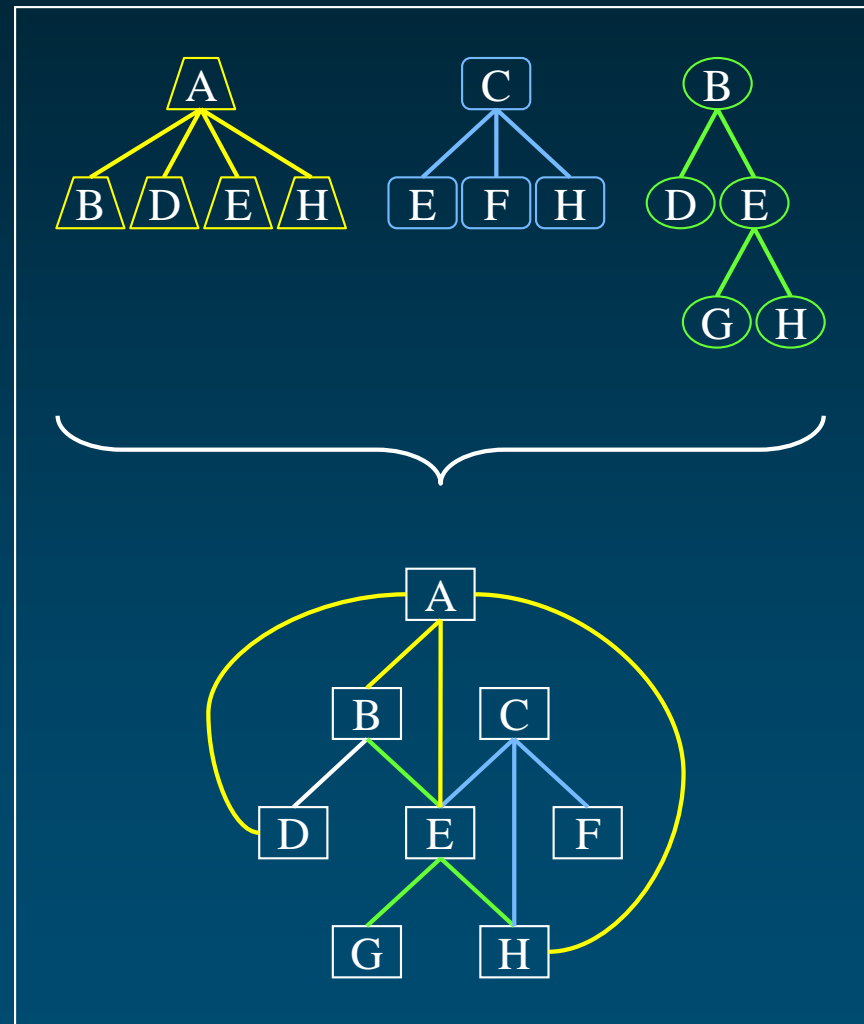
L0000002

C0000001



Organize concepts

- ◆ Inter-concept relationships: hierarchies from the source vocabularies
- ◆ Redundancy: multiple paths
- ◆ One *graph* instead of multiple *trees* (multiple inheritance)



Semantic Types

Anatomical Structure

Fully Formed Anatomical Structure

Embryonic Structure

Body Part, Organ or Organ Component

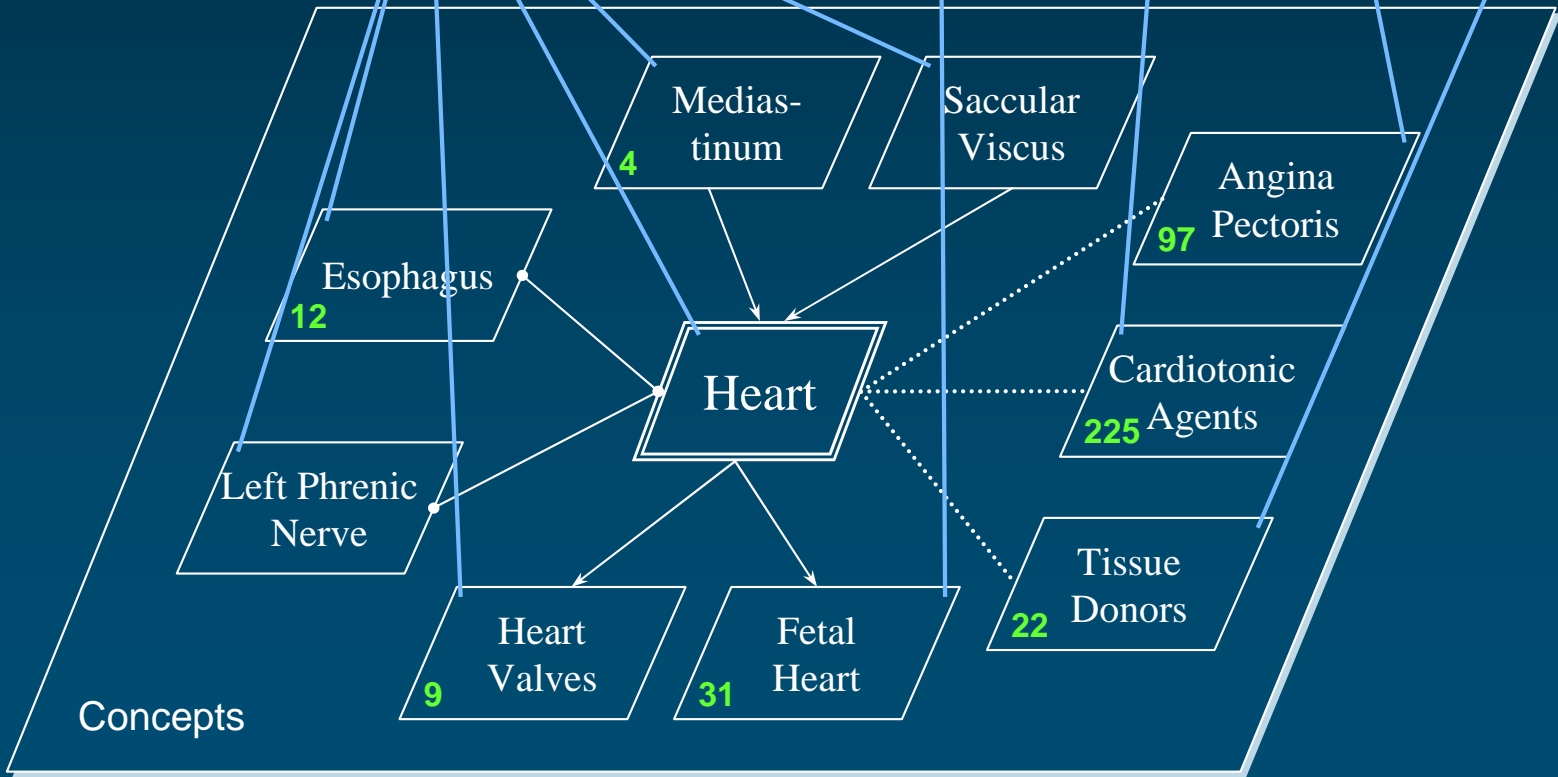
Disease or Syndrome

Pharmacologic Substance

Population Group

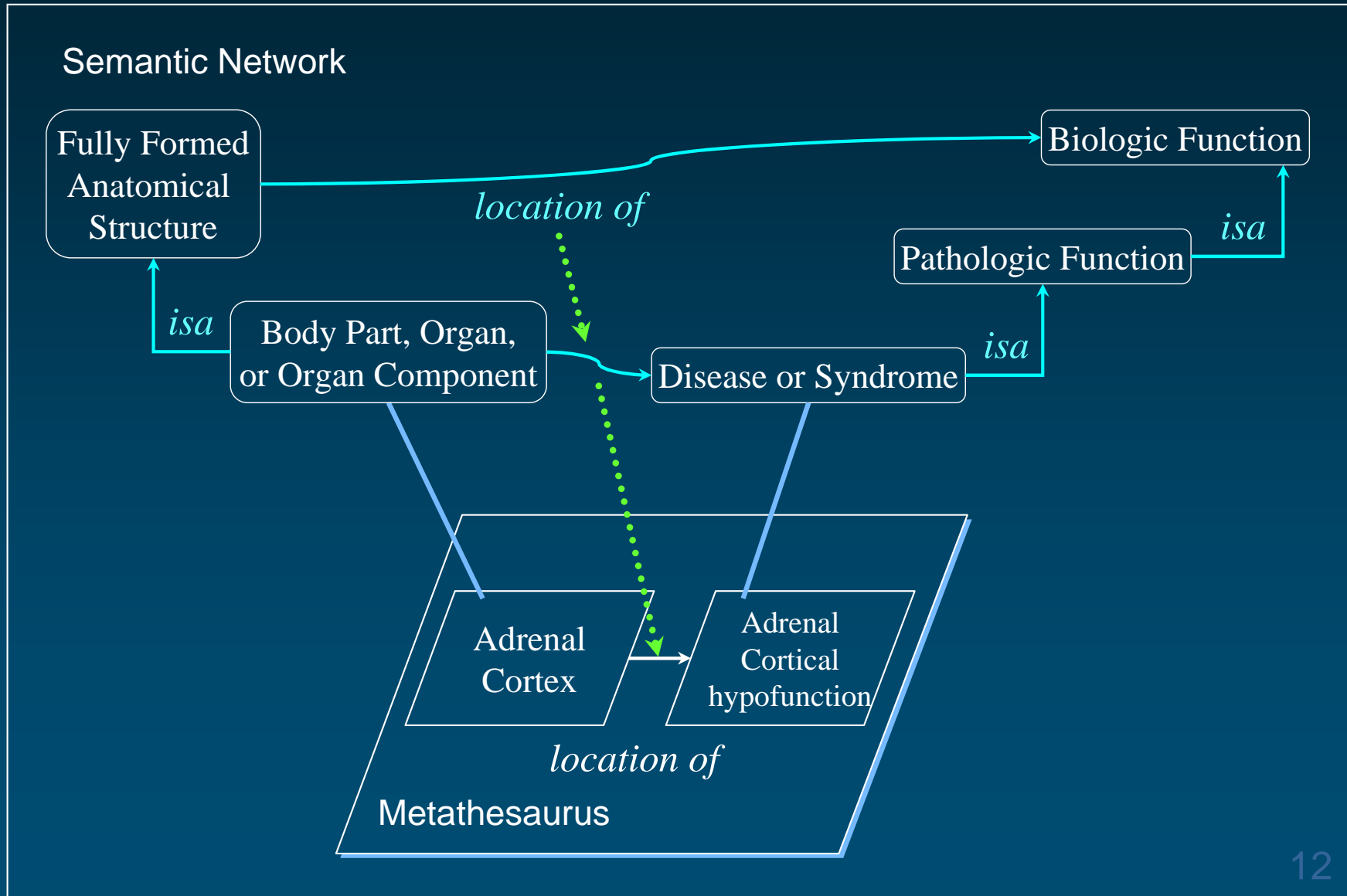
Semantic Network

Metathesaurus



Concepts

Relationships can inherit semantics



Immunology in the UMLS

◆ Semantic Network

- Immunologic Factor

◆ Metathesaurus

- Immunologic factors (C0021054)
- Immunity (C0020964)
- Autoimmune diseases (C0004364)
- Vaccines (C0042210)

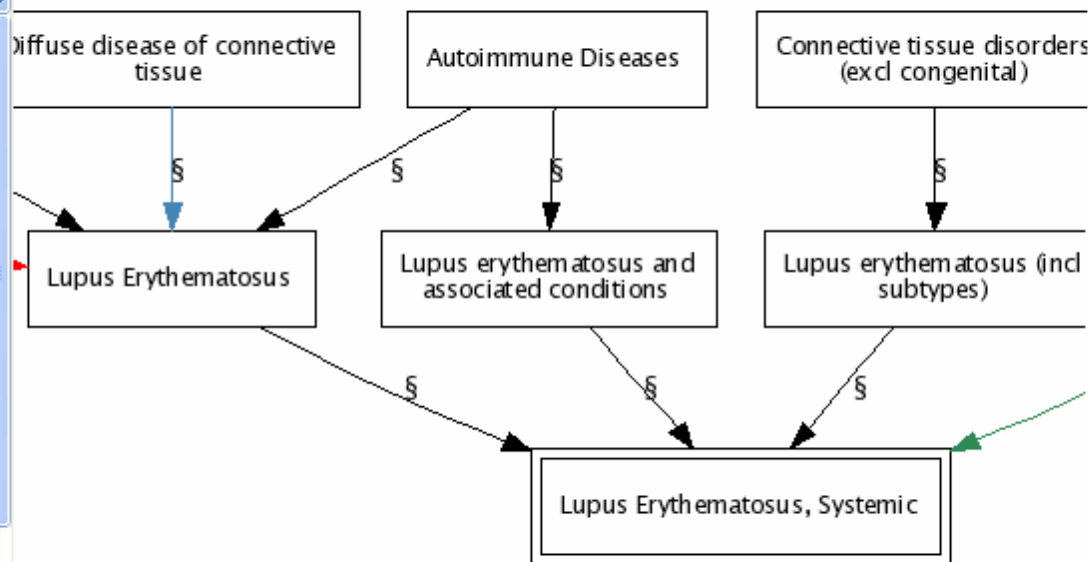
Siblings

Concepts & Ideas

- None of the above ☐
- Unsure of diagnosis ☐

Disorders

- Inflammatory Monoarthritis ☐
- Lupoid hepatic cirrhosis ☐
- Lupus endocarditis ☐
- Lupus Erythematosus, Cutaneous ☐
- Lupus erythematosus, localized, discoid ☐
- Lupus erythematosus, subacute ☐
- Lupus-like syndrome ☐



Descendants (first generation)

Disorders

- Acute systemic lupus erythematosus ☐
- Bullous systemic lupus erythematosus ☐
- dementia; systemic lupus erythematosus (etiology) ☐
- Drug-induced lupus erythematosus ☐
- Fulminating systemic lupus erythematosus ☐
- LE syndrome aggravated ☐
- LE test abnormality ☐

Other Related Concepts

Anatomy

- Connective Tissue ☐
- LE Cells ☐

Chemicals & Drugs

- Aurothioglucose ☐
- Aurothioglucose

- Anti-Arrhythmia Agents [8] ☐
- Anti-Bacterial Agents [30] ☐
- Antibodies [39] ☐
- Antibodies, Anticardiolipin [12] ☐
- Antibodies, Anti-Idiotypic [34] ☐
- Antibodies, Antinuclear [407] ☐
- Antibodies, Bacterial [9] ☐
- Antibodies, Catalytic [81] ☐

Lupus Erythematosus, Systemic

Start again Apply new parameters

Restrict to vocabulary:

Highlight vocabulary:

- ### Similar Concepts
- Lupus Nephritis ☐
 - SLE ☐

- ### Closest MeSH Terms
- #### Main Headings

Semantic Web

Web vs. Semantic web

◆ Web

- For humans
- Cosmetic markup
- Knowledge embedded in text

- Heterogeneity

- Limited interoperability

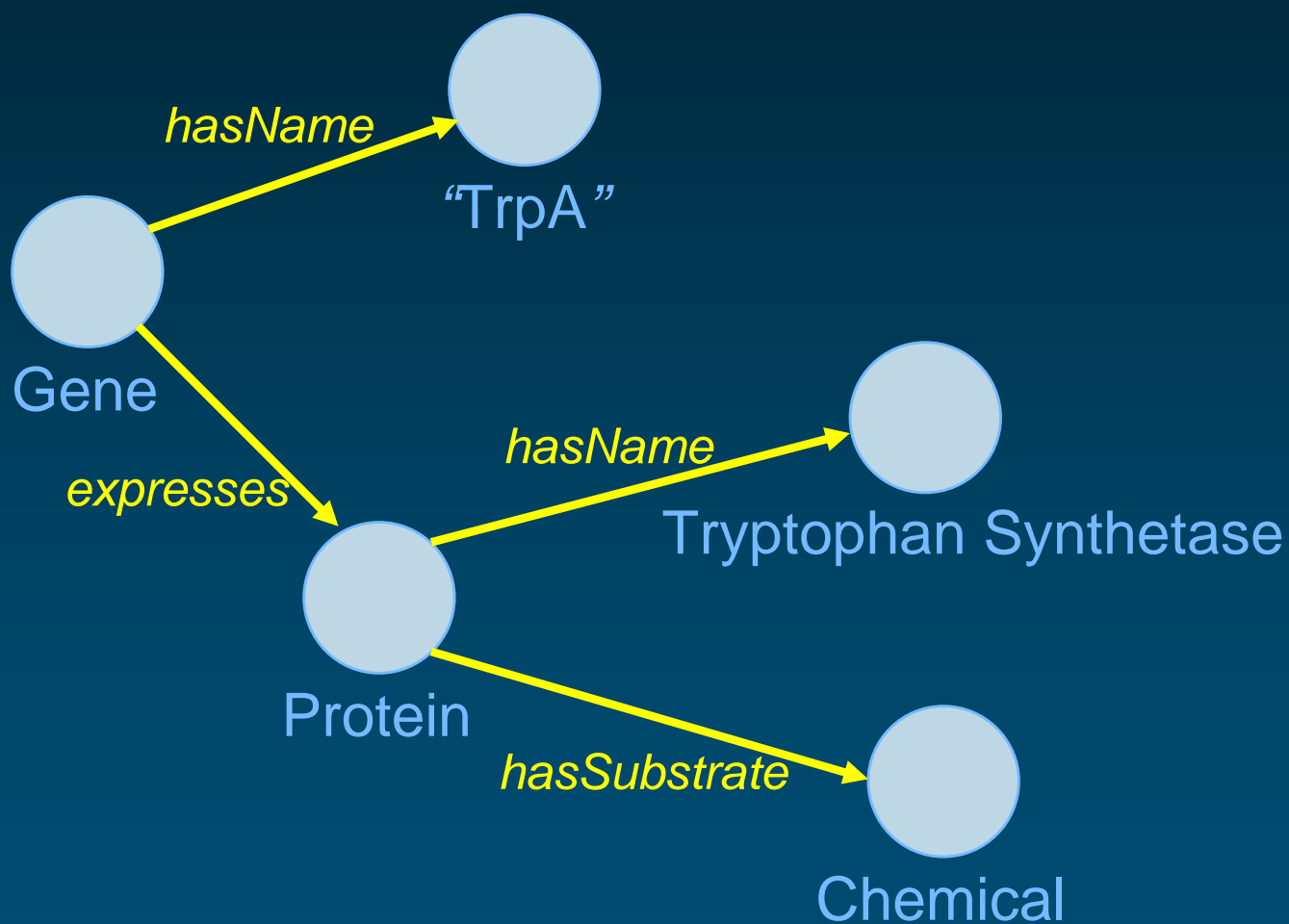
◆ Semantic web

- For humans and machines
- Semantic markup
- Knowledge represented explicitly
 - Semantic annotation (RDF)
 - Ontologies
- Integrates heterogeneous resources
- Interoperability

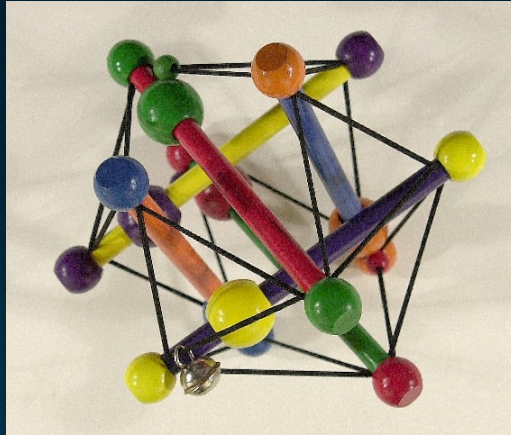
Knowledge in the semantic web

- ◆ Uniform Resource Identifier (URI)
 - Identify entities and relations
 - E.g., GO or UMLS identifier
- ◆ Resource Description Framework (RDF)
 - Triples (entity_1 , relationship, entity_2)
 - Common syntax
 - Semantic bus for applications

RDF triples



Issues and Challenges



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

UMLS References

◆ UMLS

umlsinfo.nlm.nih.gov

◆ UMLS browsers

(free, but UMLS license required)

- Knowledge Source Server: umlsks.nlm.nih.gov
- Semantic Navigator:
<http://mor.nlm.nih.gov/perl/semnav.pl>
- RRF browser
(standalone application distributed with the UMLS)



UMLS References

◆ Recent overviews

- Bodenreider O. (2004). The Unified Medical Language System (UMLS): Integrating biomedical terminology. *Nucleic Acids Research*; D267-D270.
- Nelson, S. J., Powell, T. & Humphreys, B. L. (2002). The Unified Medical Language System (UMLS) Project. In: Kent, Allen; Hall, Carolyn M., editors. *Encyclopedia of Library and Information Science*. New York: Marcel Dekker. p.369-378.

UMLS References

◆ UMLS as a research project

- Lindberg, D. A., Humphreys, B. L., & McCray, A. T. (1993). The Unified Medical Language System. *Methods Inf Med, 32*(4), 281-91.
- Humphreys, B. L., Lindberg, D. A., Schoolman, H. M., & Barnett, G. O. (1998). The Unified Medical Language System: an informatics research collaboration. *J Am Med Inform Assoc, 5*(1), 1-11.

UMLS References

◆ Technical papers

- McCray, A. T., & Nelson, S. J. (1995). The representation of meaning in the UMLS. *Methods Inf Med*, 34(1-2), 193-201.
- Bodenreider O. & McCray A. T. (2003). Exploring semantic groups through visual approaches. *Journal of Biomedical Informatics*, 36(6), 414-432.

Semantic Web References

- ◆ World Wide Web Consortium (W3C)
 - <http://www.w3.org/>
- ◆ W3C Health Care and Life Sciences Interest Group
 - <http://www.w3.org/2001/sw/hcls/>