Medinfo 2007
Tutorial S009
August 19, 2007

The Unified Medical Language System
What is it and how to use it?

Olivier Bodenreider, MD, PhD

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

Outline

- What is the UMLS?
  - Introduction
  - Overview through an example
  - The three UMLS Knowledge Sources
- How to use the UMLS?
  - Obtaining a license
  - Remote access
  - Local installation and customization
  - A UMLS-based algorithm
  - Benefits and limitations

Part I

What is the UMLS?

Outline

- Part I: *What is the UMLS?*
  - Introduction
  - Overview through an example
  - The three UMLS Knowledge Sources
    - UMLS Metathesaurus
    - UMLS Semantic Network
    - SPECIALIST Lexicon and lexical tools

Part I
What is the UMLS?

(1) Introduction

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"
- Complementary to IAIMS

(Integrated Academic Information Management Systems)

«[...] 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.»

The UMLS in practice

- Database
  - Series of relational files
- Interfaces
  - Web interface: Knowledge Source Server (UMLSKS)
  - Application programming interfaces (Java and XML-based)
- Applications
  - lvg (lexical programs)
  - MetamorphoSys (installation and customization)
  - RRF browser (browsing subsets)

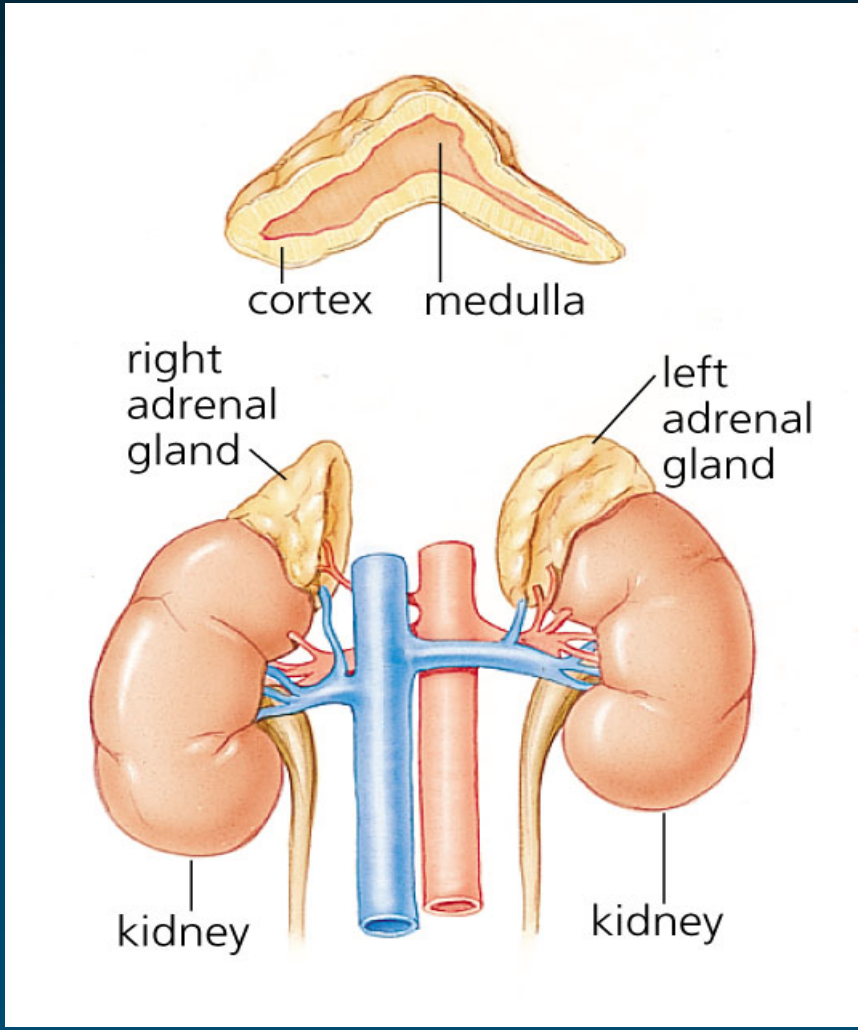
The UMLS is *not* an end-user application

Part I
What is the UMLS?

(2) Overview through an example

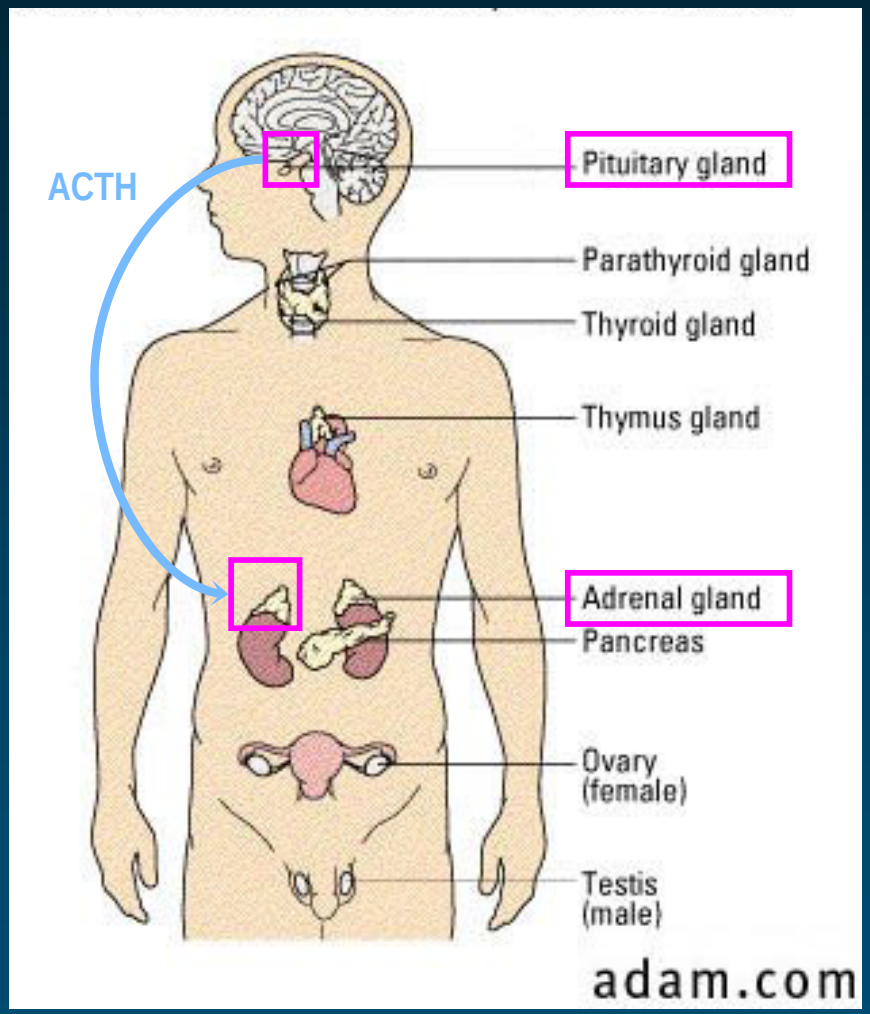
Addison’s disease

- Addison's disease is a rare endocrine disorder
- Addison's disease occurs when the adrenal glands do not produce enough of the hormone cortisol
- For this reason, the disease is sometimes called chronic adrenal insufficiency, or hypocortisolism



Adrenal insufficiency Clinical variants

- Primary / Secondary
  - Primary: lesion of the adrenal glands themselves
  - Secondary: inadequate secretion of ACTH by the pituitary gland
- Acute / Chronic
- Isolated / Polyendocrine deficiency syndrome



Addison's disease: Symptoms

- Fatigue
- Weakness
- Low blood pressure
- Pigmentation of the skin (exposed and non-exposed parts of the body)
- …

AD in medical vocabularies

- Synonyms: different terms
  - Addisonian syndrome — eponym
  - Bronzed disease — symptoms
  - Melasma addisonii — symptoms
  - Asthenia pigmentosa — symptoms
  - Primary adrenal deficiency — clinical variants
  - Primary adrenal insufficiency — clinical variants
  - Primary adrenocortical insufficiency — clinical variants
  - Chronic adrenocortical insufficiency — clinical variants
- Contexts: different hierarchies

Organize terms

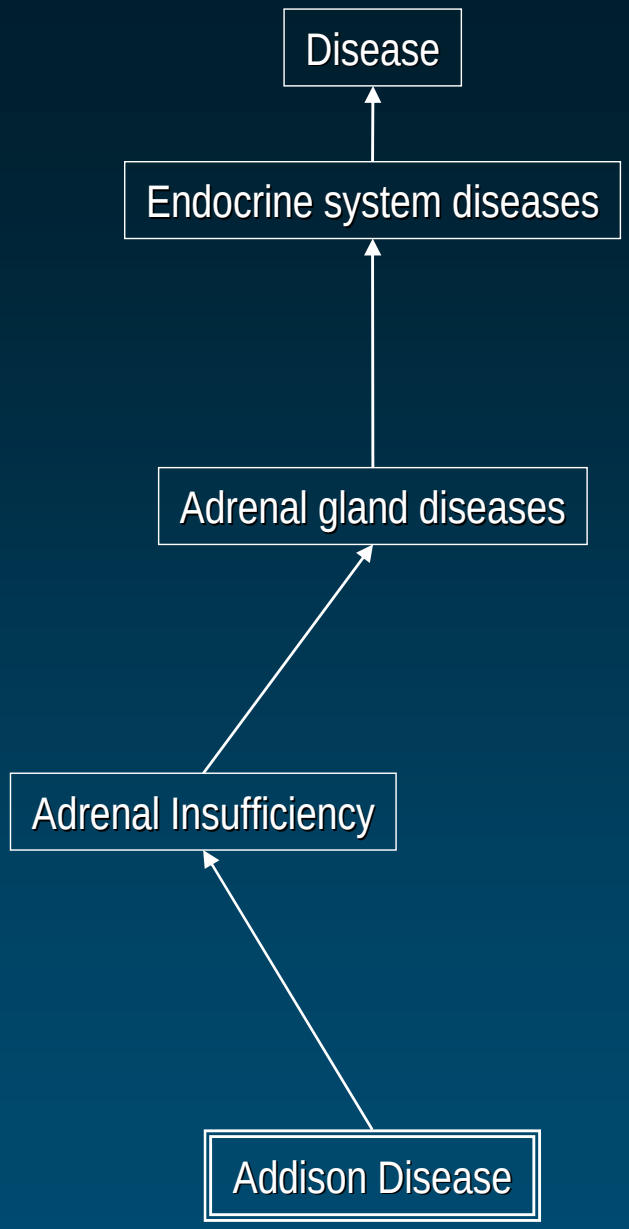
- Synonymous terms clustered into a concept
- Preferred term
- Unique identifier (CUI)

Addison Disease	MeSH	D000224
Primary hypoadrenalism	MedDRA	10036696
Primary adrenocortical insufficiency	ICD-10	E27.1
Addison's disease (disorder)	SNOMED CT	363732003

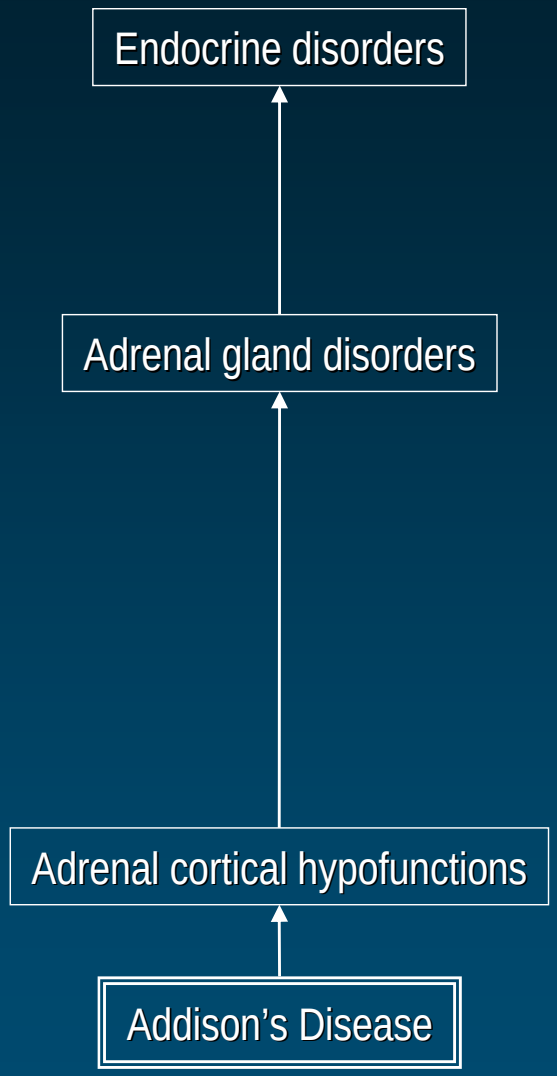
C0001403

Addison's disease

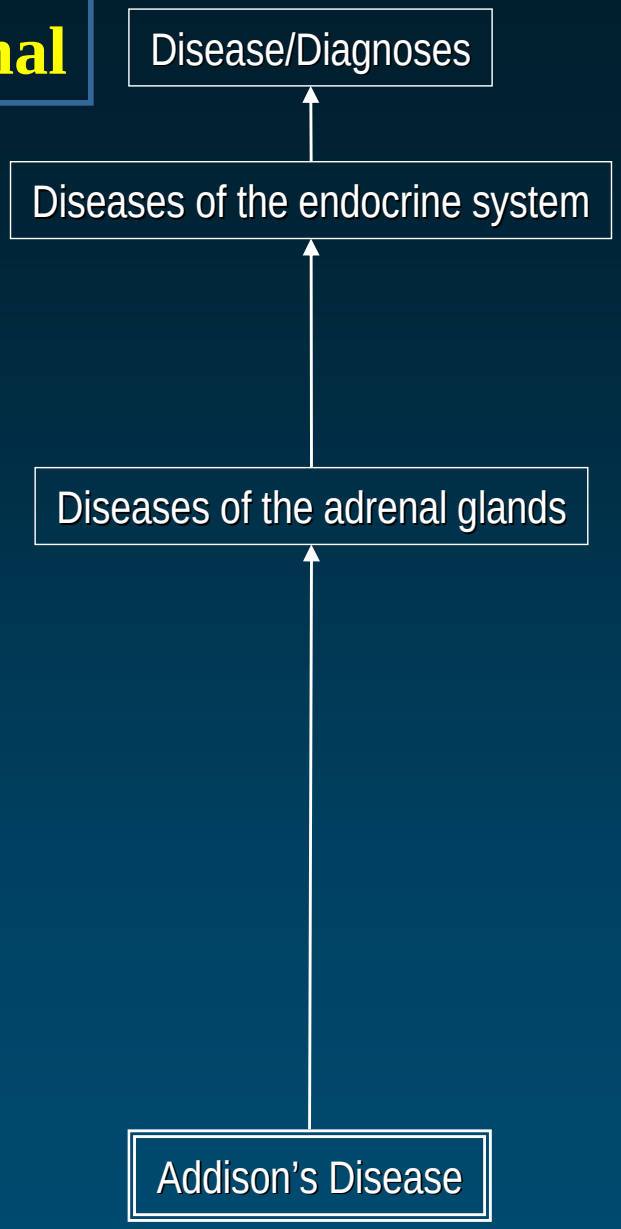
MeSH


Disease
Endocrine system diseases
Adrenal gland diseases
Adrenal Insufficiency
Addison Disease

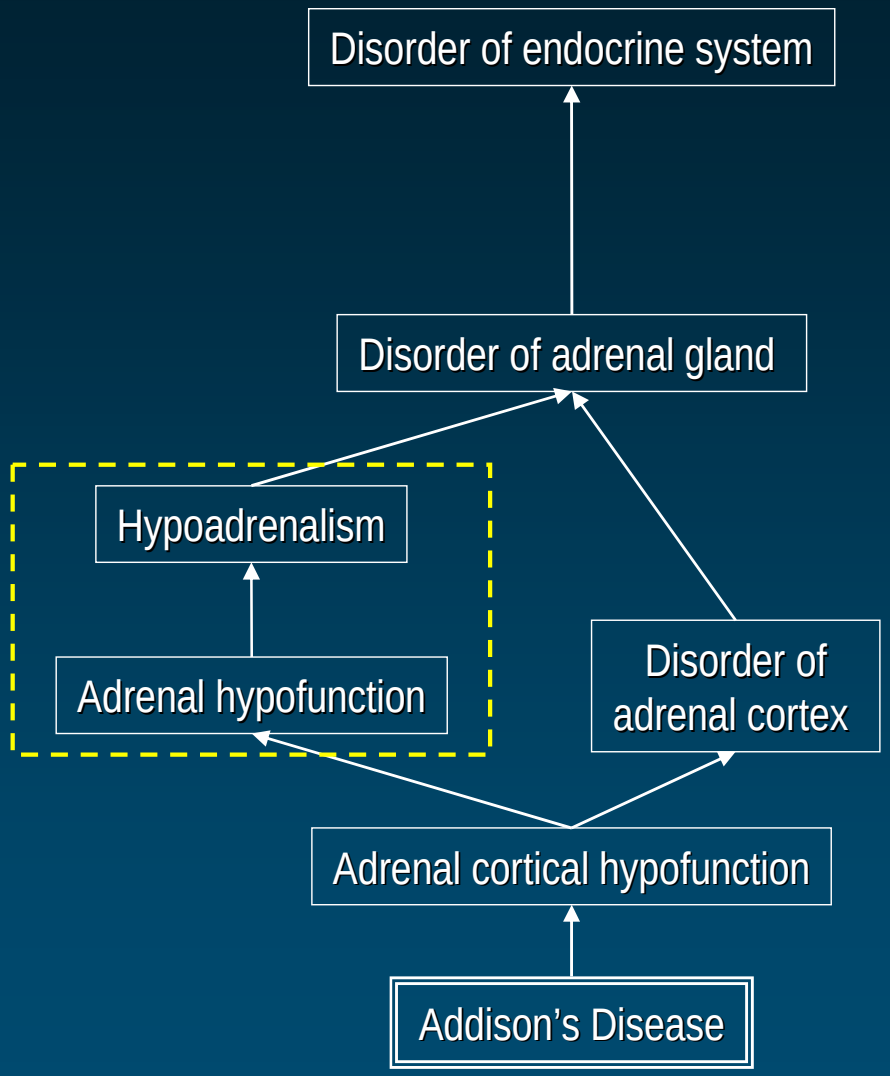
MedDRA


Endocrine disorders
Adrenal gland disorders
Adrenal cortical hypofunctions
Addison's Disease

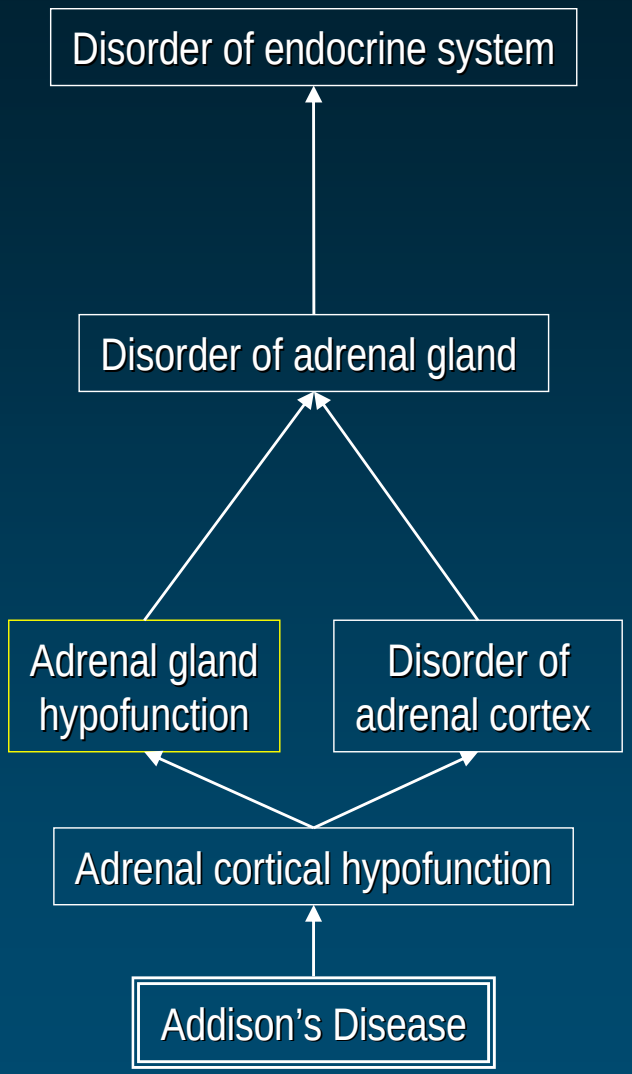
SNOMED International



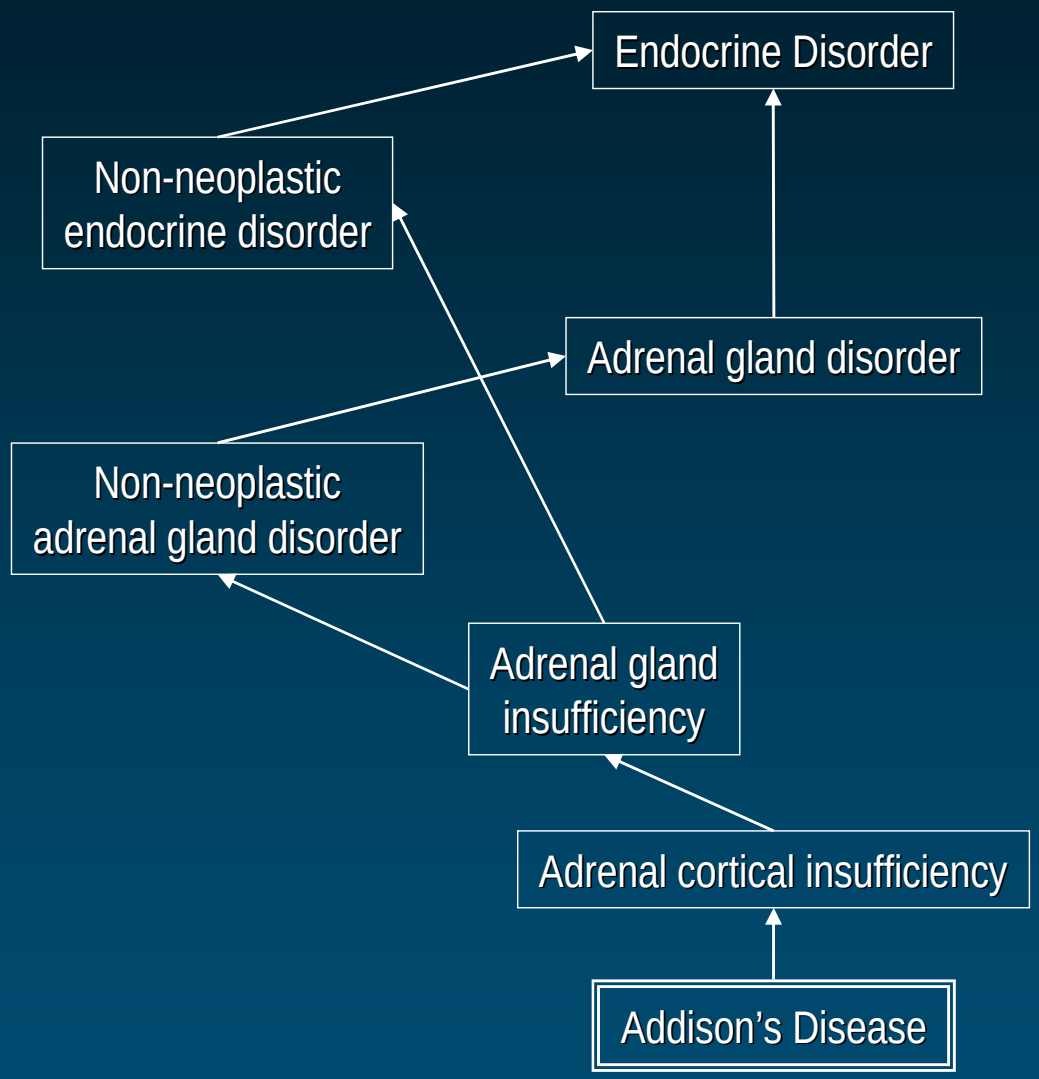
SNOMED CT (native)



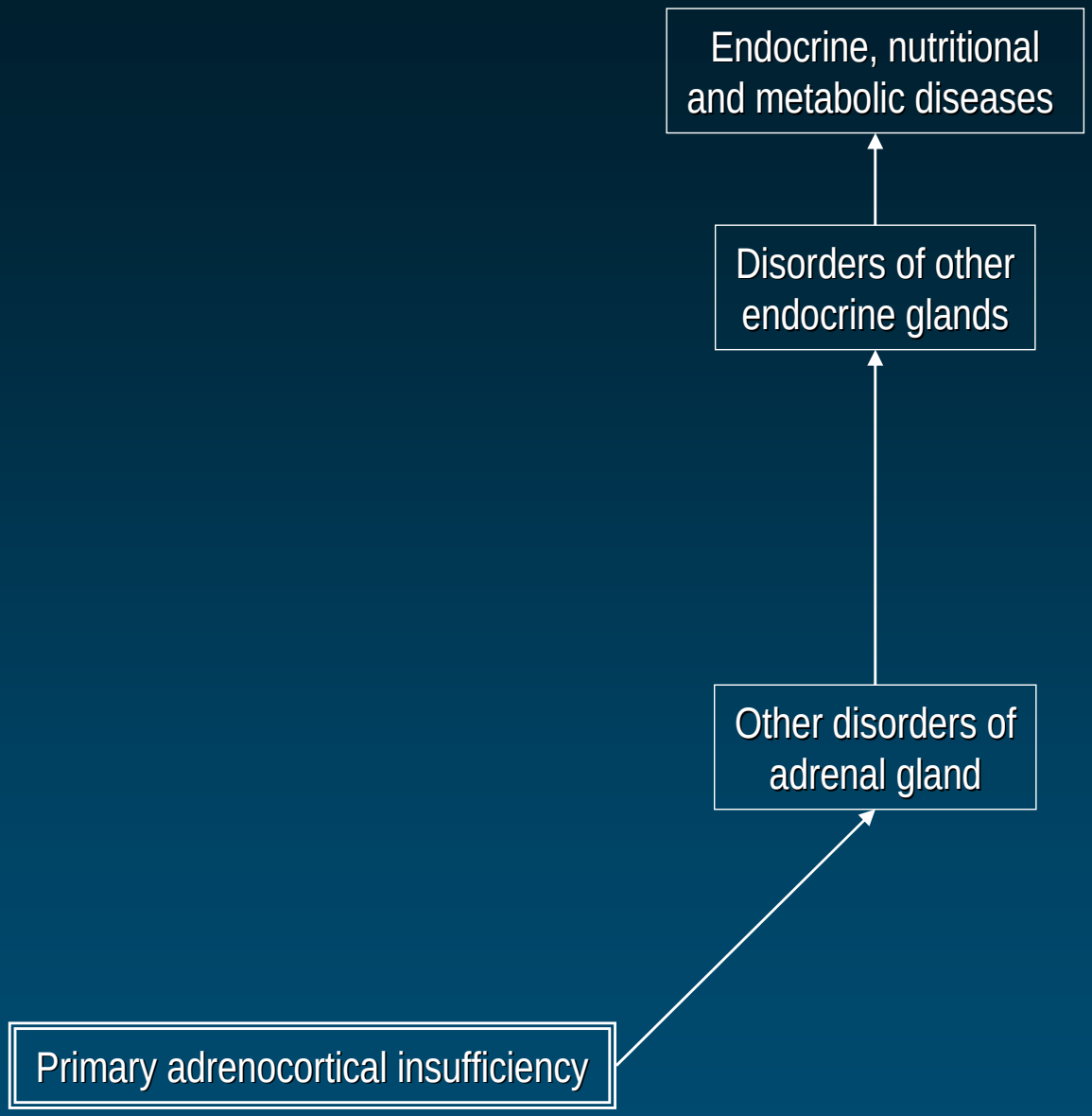
SNOMED CT (UMLS view)



NCI Thesaurus

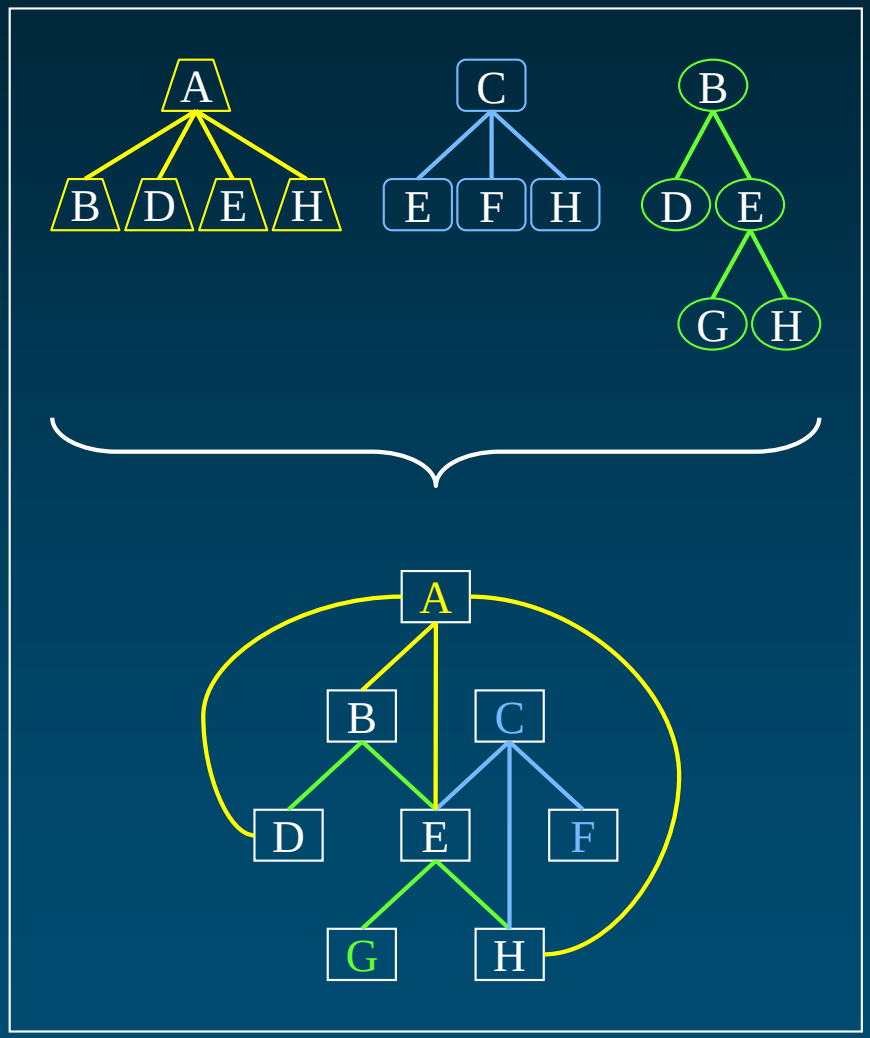


ICD-10

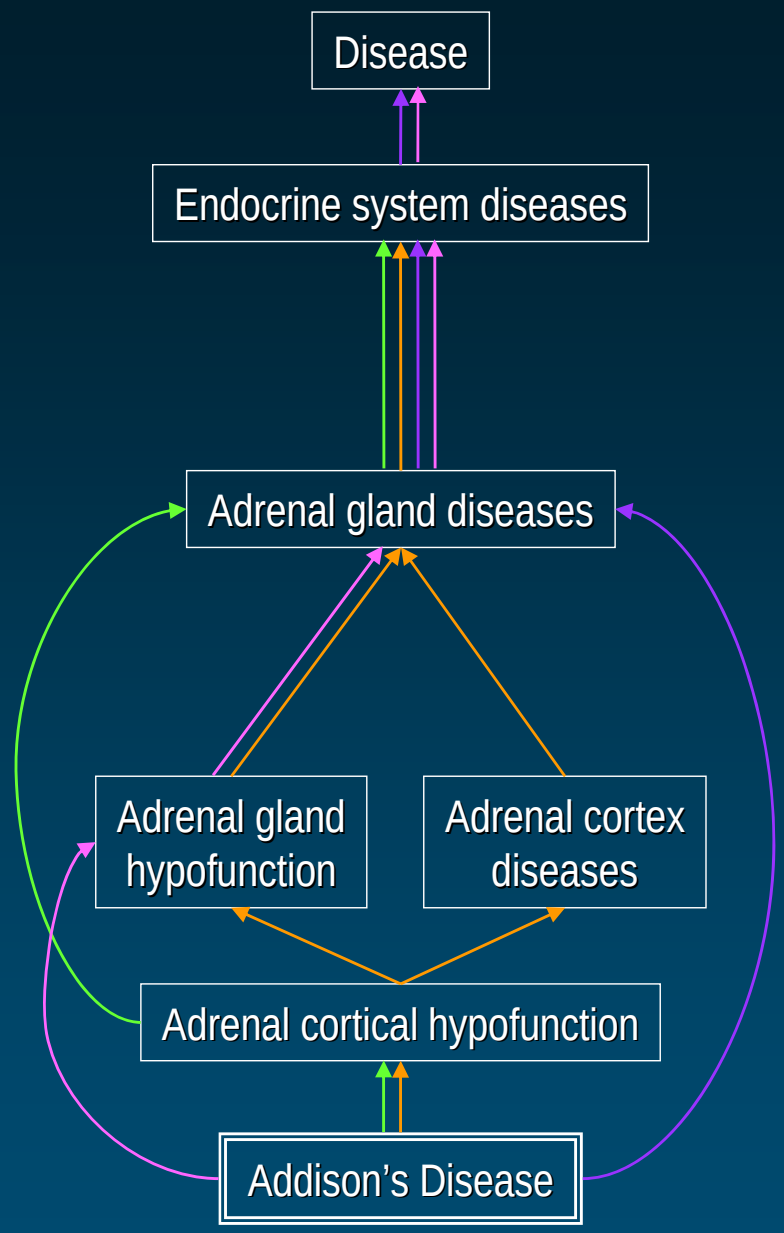

Endocrine, nutritional and metabolic diseases
Disorders of other endocrine glands
Other disorders of adrenal gland
Primary adrenocortical insufficiency

Organize concepts

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

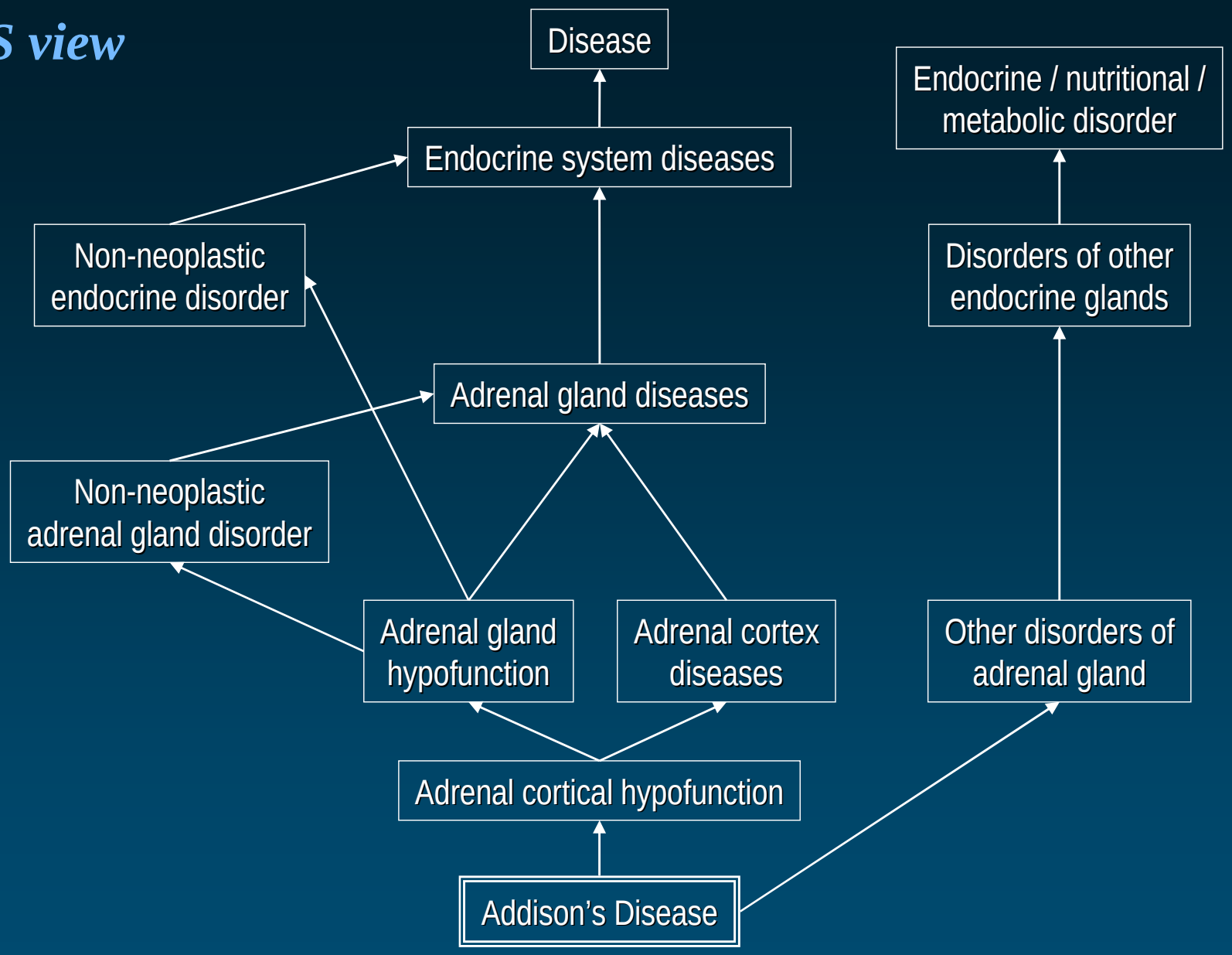


organize concepts

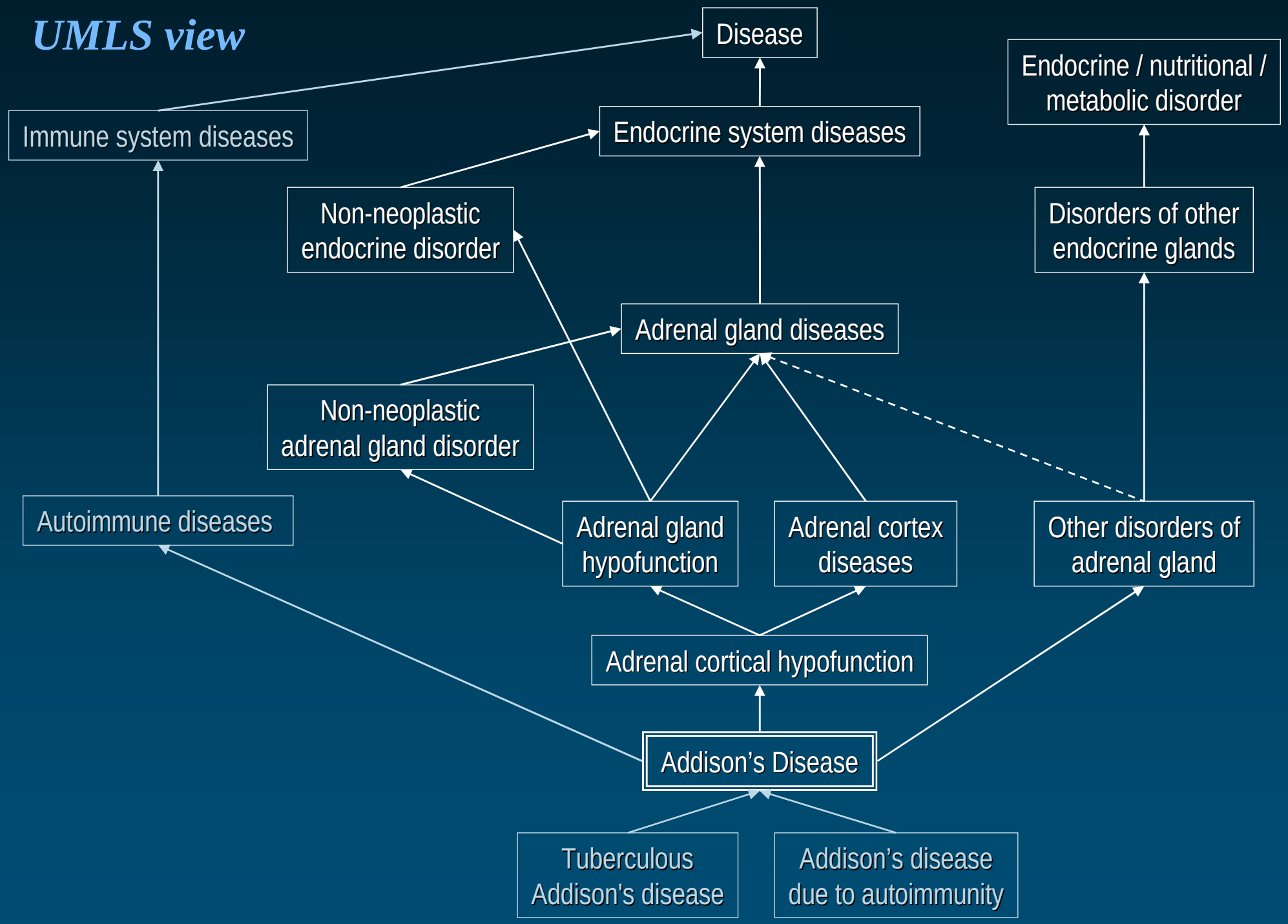


SNOMED CT
SNOMED Intl
MeSH
MedDRA

UMLS view


Disease
Endocrine / nutritional / metabolic disorder
Endocrine system diseases
Non-neoplastic endocrine disorder
Disorders of other endocrine glands
Adrenal gland diseases
Non-neoplastic adrenal gland disorder
Adrenal gland hypofunction
Adrenal cortex diseases
Other disorders of adrenal gland
Adrenal cortical hypofunction
Addison's Disease

UMLS view

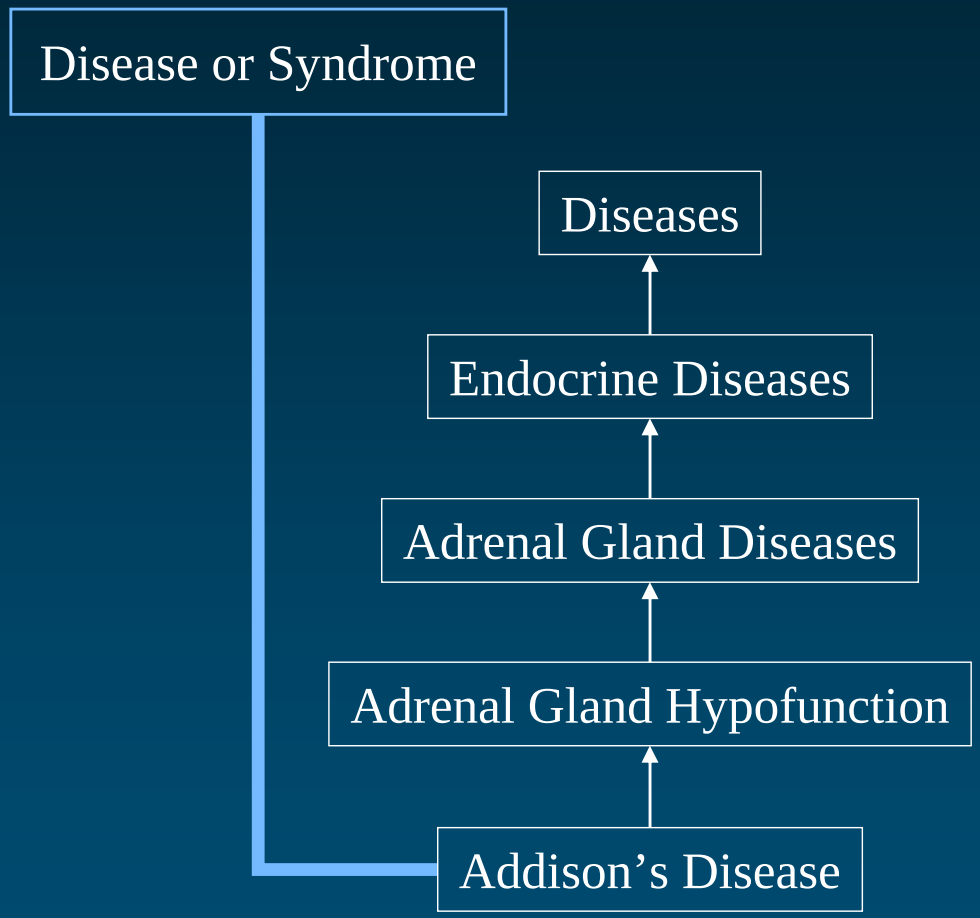

Disease
Endocrine / nutritional / metabolic disorder
Immune system diseases
Endocrine system diseases
Non-neoplastic endocrine disorder
Disorders of other endocrine glands
Adrenal gland diseases
Non-neoplastic adrenal gland disorder
Autoimmune diseases
Adrenal gland hypofunction
Adrenal cortex diseases
Other disorders of adrenal gland
Adrenal cortical hypofunction
Addison's Disease
Tuberculous Addison's disease
Addison's disease due to autoimmunity

Relate to other concepts

- Additional hierarchical relationships
  - link to other trees
  - make relationships explicit
- Non-hierarchical relationships
- Co-occurring concepts
- Mapping relationships

Categorize concepts

- High-level categories (semantic types)
- Assigned by the Metathesaurus editors
- Independently of the hierarchies in which these concepts are located



How do they do that?

- Lexical knowledge
- Semantic pre-processing
- UMLS editors

Lexical knowledge

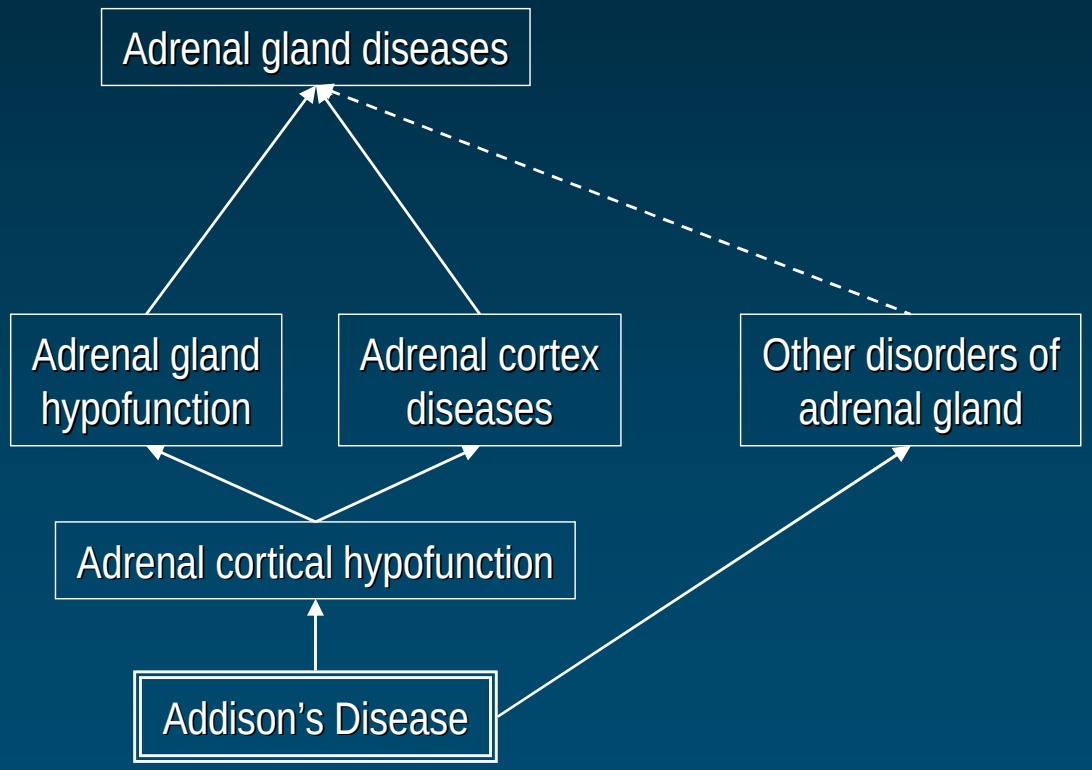
Adrenal gland diseases
Adrenal disorder
Disorder of adrenal gland
Diseases of the adrenal glands
C0001621

Semantic pre-processing

- Metadata in the source vocabularies

- Tentative categorization
- Positive (or negative) evidence for tentative synonymy relations based on lexical features

Additional knowledge: UMLS editors



UMLS Summary

- Synonymous terms clustered into concepts
- Unique identifier

- Finer granularity
- Broader scope
- Additional hierarchical relationships
- Semantic categorization

Part I
What is the UMLS?

(3) UMLS Knowledge Sources

Unified Medical Language System

- SPECIALIST Lexicon
  - 360,000 lexical items
  - Part of speech and variant information
- Metathesaurus
  - 6M names from over 100 terminologies
  - 1.5M concepts
  - 8M relations
- Semantic Network
  - 135 high-level categories
  - 7000 relations among them

Lexical resources

Terminological resources

Ontological resources

UMLS Metathesaurus

Metathesaurus Basic organization

- Concepts
  - Synonymous terms are clustered into a concept
  - Properties are attached to concepts, e.g.,
    - Unique identifier
    - Definition
- Relations
  - Concepts are related to other concepts
  - Properties are attached to relations, e.g.,
    - Type of relationship
    - Source

Source Vocabularies

(2007AB)

- 143 source vocabularies
  - 17 languages
- Broad coverage of biomedicine
  - 5.9M names
  - 1.4M concepts
  - 8M relations
- Common presentation

Biomedical terminologies

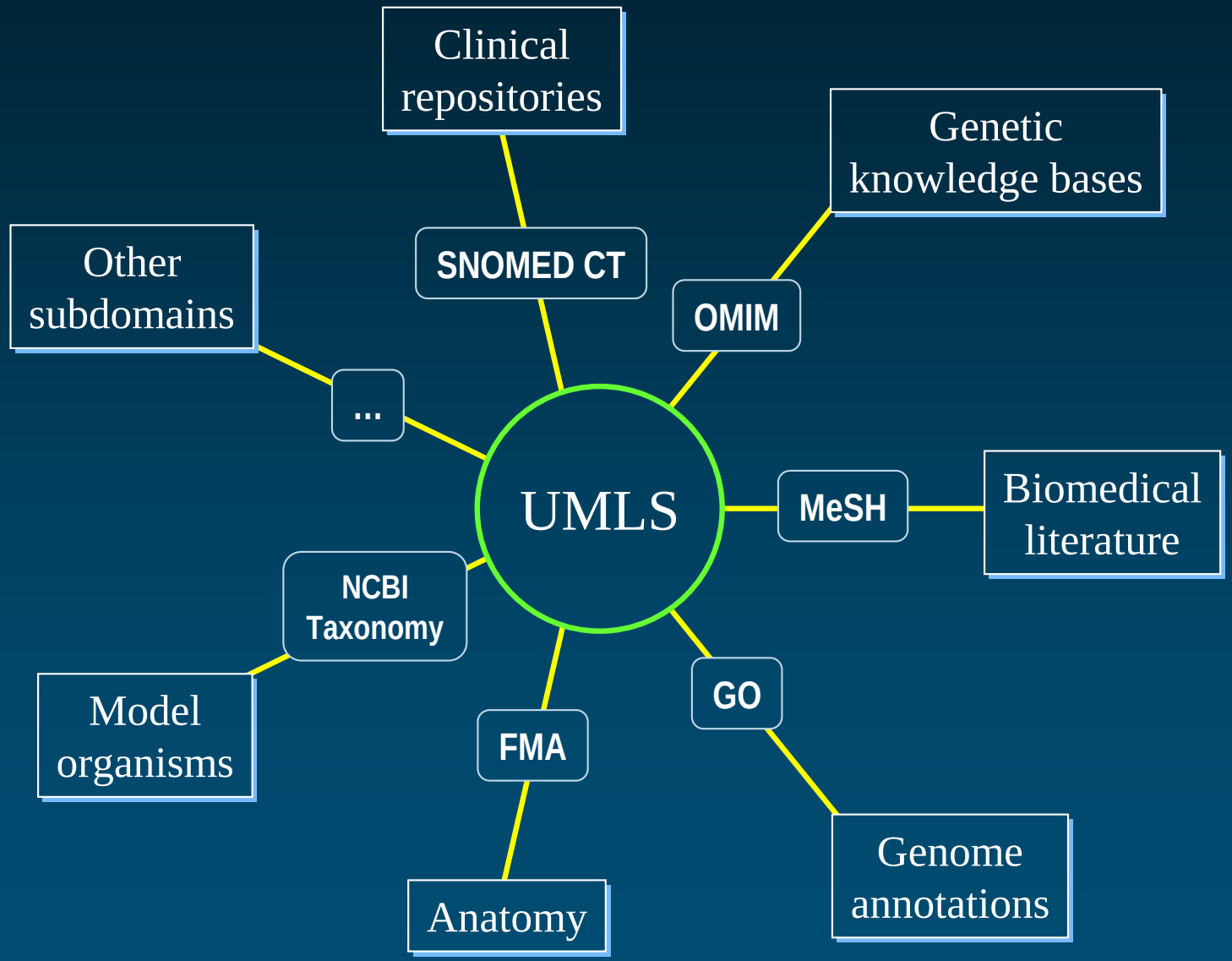
- General vocabularies
  - anatomy (UWDA, Neuronames)
  - drugs (RxNorm, First DataBank, Micromedex)
  - medical devices (UMD, SPN)
- Several perspectives
  - clinical terms (SNOMED CT)
  - information sciences (MeSH, CRISP)
  - administrative terminologies (ICD-9-CM, CPT-4)
  - data exchange terminologies (HL7, LOINC)

Biomedical terminologies (cont'd)

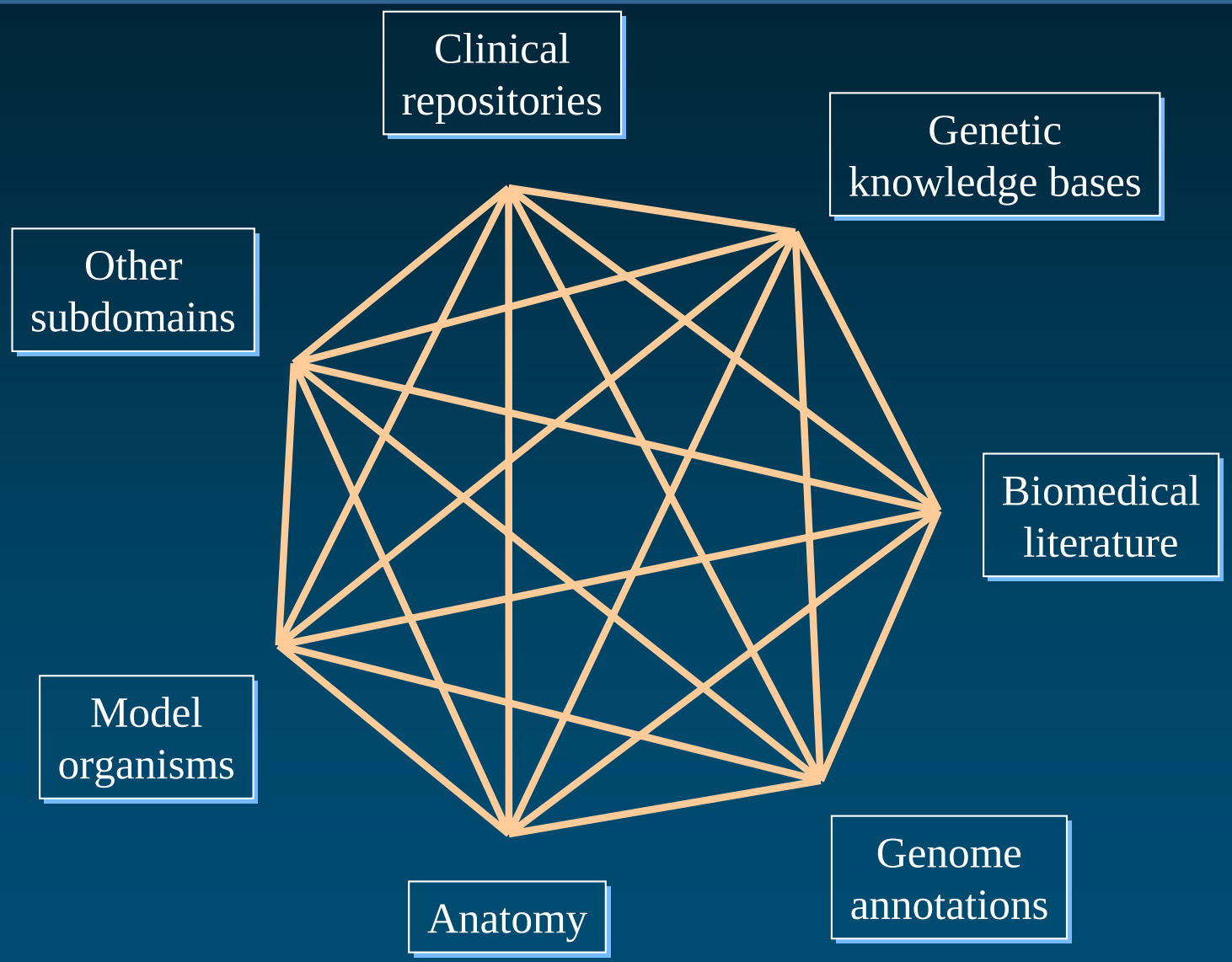
- Specialized vocabularies
  - nursing (NIC, NOC, NANDA, Omaha, PCDS)
  - dentistry (CDT)
  - oncology (PDQ)
  - psychiatry (DSM, APA)
  - adverse reactions (COSTART, WHO ART)
  - primary care (ICPC)
- Terminology of knowledge bases (AI/Rheum, DXplain, QMR)

The UMLS serves as a vehicle for the regulatory standards (HIPAA, CHI)

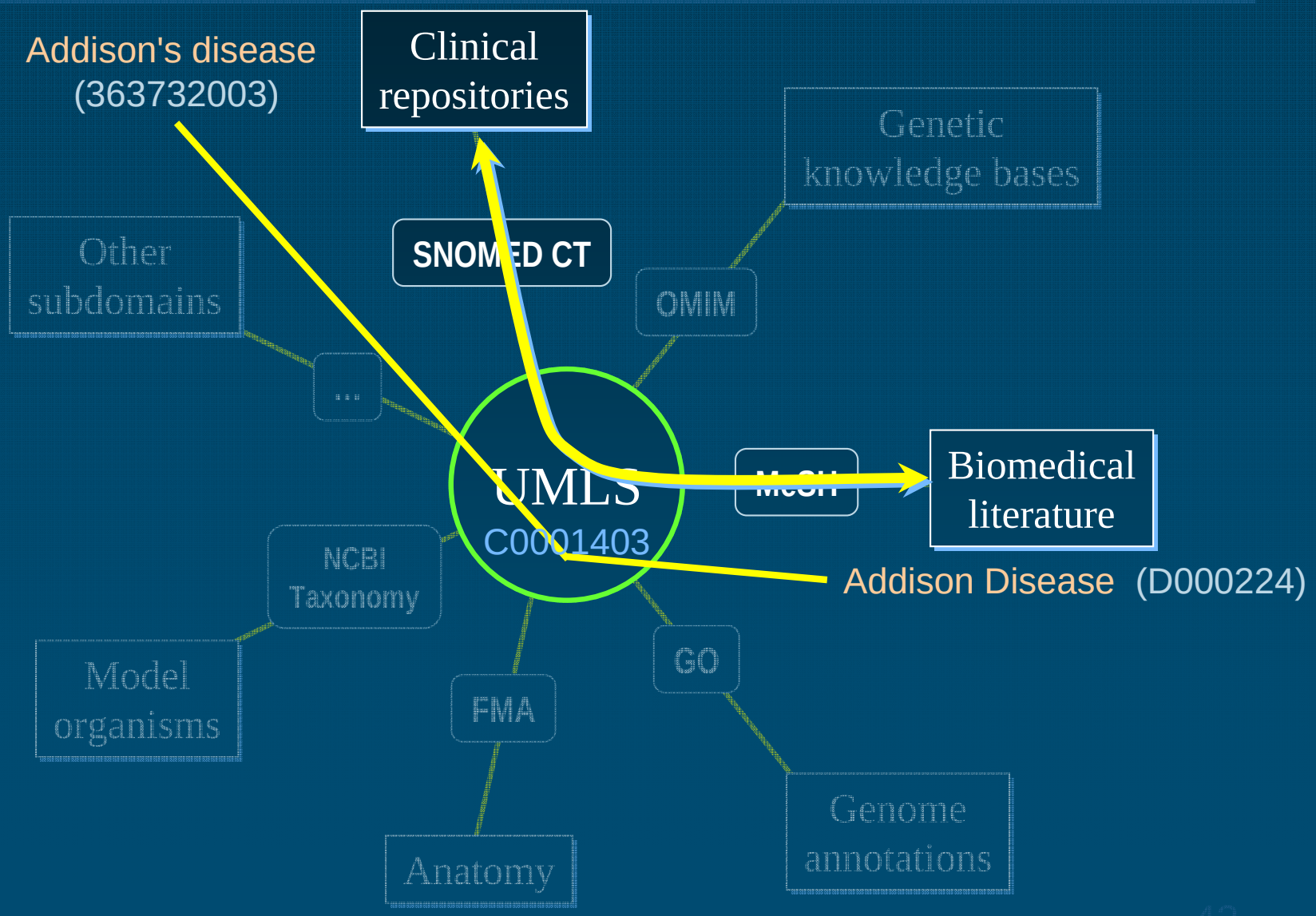
Integrating subdomains



Integrating subdomains



Trans-namespace integration



Addison's Disease: Concept

Disease or Syndrome

Addison's Disease

ADRENAL INSUFFICIENCY (ADDISON'S DISEASE)
ADRENOCORTICAL INSUFFICIENCY, PRIMARY FAILURE
Hypoadrenalisms, Primary
Melasma addisonii
Primary adrenal deficiency
Asthenia pigmentosa
Bronzed disease
Insufficiency, adrenal primary
Primary adrenocortical insufficiency
Addison's, disease

Maladie d'Addison - French
Addison-Krankheit - German
Morbo di Addison - Italian
Doença de Addison - Portuguese
АДДИСОНОВА БОЛЕЗНЬ - Russian
アジソン病 - Japanese

SNOMED CT
SNOMED Intl
MeSH
MedDRA
…

C0001403

An adrenal disease characterized by the progressive destruction of the adrenal cortex, resulting in insufficient production of aldosterone and hydrocortisone. Clinical symptoms include anorexia; nausea; weight loss; muscle ewakness; and hyperpigmentation of the skin due to increase in circulating levels of ACTH precursor hormone which stimulates melanocytes.

NLM

Metathesaurus Concepts

(2007AB)

- Concept (~ 1.4M) CUI
  - Set of synonymous concept names
- Term (~ 5.3 M) LUI
  - Set of normalized names
- String (~ 5.9M) SUI
  - Distinct concept name
- Atom (~ 7.2M) AUI
  - Concept name in a given source

A0066000 Headache (MeSH)
A0065992 Headache (ICD-10)
S0046854

A0066007 Headaches (MedDRA)
A12003304 Headaches (OMIM)
S0046855

L0018681

A0540936 Cephalodynia (MeSH)
S0475647

L0380797

C0018681

Cluster of synonymous terms

Concept
C0001403

Term L0001403	**S0354372** *Addison's disease* **S0010794** Addison's Disease **S0010792** Addison Disease **S0010796** Addisons Disease **S0033587** Disease, Addison **S0469271** Addison's disease, NOS	[...]	
Term L0494940	**S5907336** *Primary Adrenocortical Insufficiency* **S5901878** Insufficiencies, Primary Adrenocortical		
Term L0494851	**S5907334** *Primary Adrenal Insufficiency* **S5924573** Adrenal Insufficiency, Primary	[...]	
Term L0585243	**S5907343** *Primary Hypoadrenalism* **S0718109** Primary hypoadrenalism	[...]	
Term L3541031	**S4115514** *primary; hypoadrenocorticism* **S4090095** hypoadrenocorticism; primary	[...]	
Term L1229627	**S1471573** *Addison-Krankheit*	GER	
Term L5345155	**S6107160** *Maladie d'Addison*	FRE	[...]

Beyond concepts Descriptor level

- In some vocabularies, the unit of information is an aggregate of concepts
  - e.g., descriptors in MeSH, for indexing purposes

MeSH Heading	Hydrocortisone	C0020268
Tree Number	D04.808.745.745.654.600	
Tree Number	D06.472.040.585.353.476	
Tree Number	D06.472.040.585.478.392	
Scope Note	The main glucocorticoid secreted by the ADRENAL CORTEX. Its synthetic counterpart is used, either as an injection or t inflammation, allergy, collagen diseases, asthma, adrenocortical deficiency, shock, and some neoplastic conditions.	
Entry Term	11-Epicortisol	C0887247
Entry Term	Cortifair	C0699401
Entry Term	Cortisol	C0020268
Entry Term	Cortril	C0699402
Entry Term	Epicortisol	C0887247
Entry Term	Hydrocortisone, (11 alpha)-Isomer	C0887247
Entry Term	Hydrocortisone, (9 beta,10 alpha,11 alpha)-Isomer	C0887246

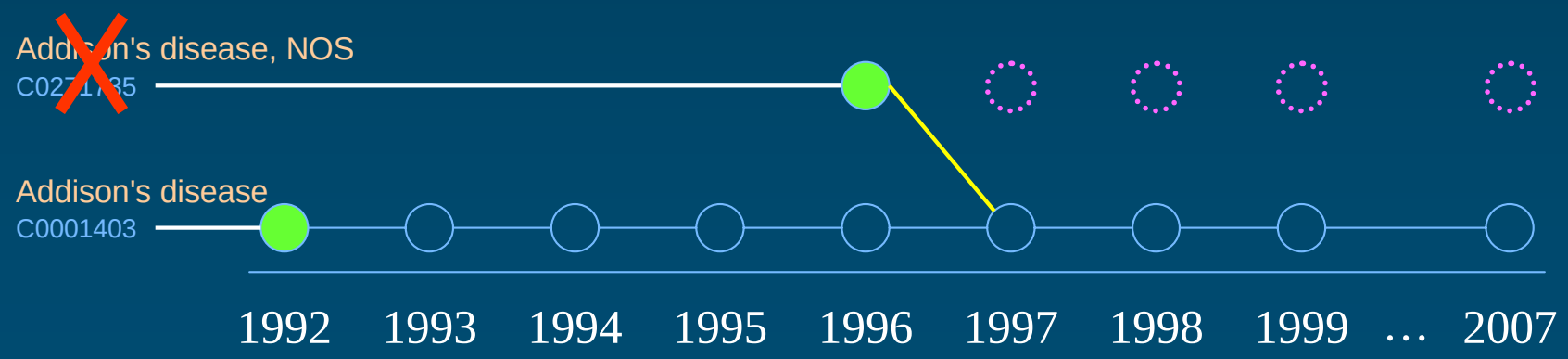
MEDICAL SUBJECT HEADINGS MeSH

Beyond concepts Descriptor level

A0070119	Hydrocortisone	(MeSH)
A0043102	Cortisol	(MeSH)
	C0020268	
A0066000	Epicortisol	(MeSH)
A0016625	11-Epicortisol	(MeSH)
A0055118	Hydrocortisone, (11 alpha)-Isomer	(MeSH)
	C0887247	
A7801724	Hydrocortisone, (9 beta,10 alpha,11 alpha)-Isomer	(MeSH)
	C0887246	
A7757592	Cortifair	(MeSH)
	C0699401	
A7757595	Cortril	(MeSH)
	C0699402	
	D006854	

Metathesaurus Evolution over time

- Concepts never die (in principle)
  - CUIs are permanent identifiers
- What happens when they do die (in reality)?
  - Concepts can merge or split
  - Resulting in new concepts and deletions



Metathesaurus Relationships

- Symbolic relations: ~8 M pairs of concepts
- Statistical relations : ~6 M pairs of concepts (co-occurring concepts)
- Mapping relations: ~150,000

- Categorization: Relationships between concepts and semantic types from the Semantic Network

Symbolic relations

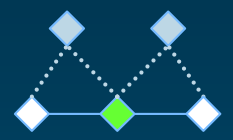
- Relation
  - Pair of “atom” identifiers
  - Type
  - Attribute (if any)
  - List of sources (for type and attribute)
- Semantics of the relationship: defined by its type [and attribute]

Source transparency: the information is recorded at the “atom” level

Symbolic relationships Type

- Hierarchical
  - Parent / Child **PAR/CHD**
  - Broader / Narrower than **RB/RN**

- Derived from hierarchies
  - Siblings (children of parents) **SIB**



- Associative
  - Other **RO**

- Various flavors of near-synonymy
  - Similar **RL**
  - Source asserted synonymy **SY**
  - Possible synonymy **RQ**

Symbolic relationships Attribute

- Hierarchical
  - isa (is-a-kind-of)
  - part-of
- Associative
  - location-of
  - caused-by
  - treats
  - …
- Cross-references (mapping)

Mapping relations

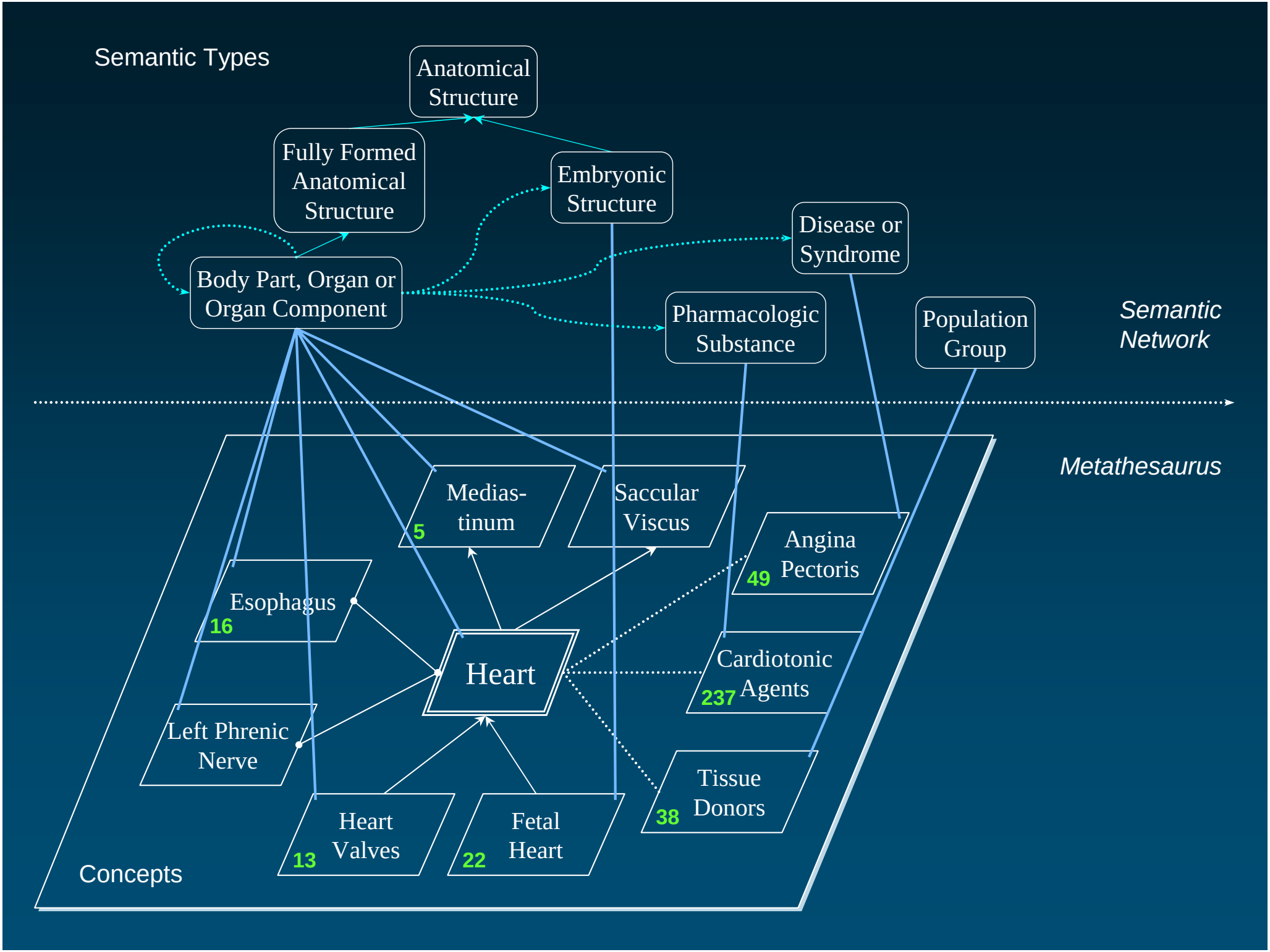
- Simple mappings
  - <atom 1> mapped_to <atom 2>
  - e.g.,
    - SNOMED CT to ICD-9-CM
- Complex mappings
  - <atom 1> mapped_to <boolean expression>
  - e.g.,
    - ICD-9-CM to MeSH (search strategies)

NB: partially redundant with relations in MRREL

Everything else

- Co-occurrence information (MRCOC)
  - Co- occurrence of MeSH descriptors in MEDLINE for the most part
- Source-specific attributes (MRSAT)
  - Legacy identifiers, external cross-references
    - SNOMED International legacy codes (SNOMED CT)
    - RxNorm to NDC
  - Concept status in a particular source (SNOMED CT)
  - Frequency of occurrence in MEDLINE (MeSH)
  - MedlinePlus URL (MeSH)
  - …

NLM


Semantic Types
Anatomical Structure
Fully Formed Anatomical Structure
Embryonic Structure
Disease or Syndrome
Body Part, Organ or Organ Component
Pharmacologic Substance
Population Group
Semantic Network
Metathesaurus
Medias-tinum
5
Saccular Viscus
Angina Pectoris
49
Esophagus
16
Heart
Cardiotonic Agents
237
Left Phrenic Nerve
Tissue Donors
38
Heart Valves
13
Fetal Heart
22
Concepts

UMLS Semantic Network

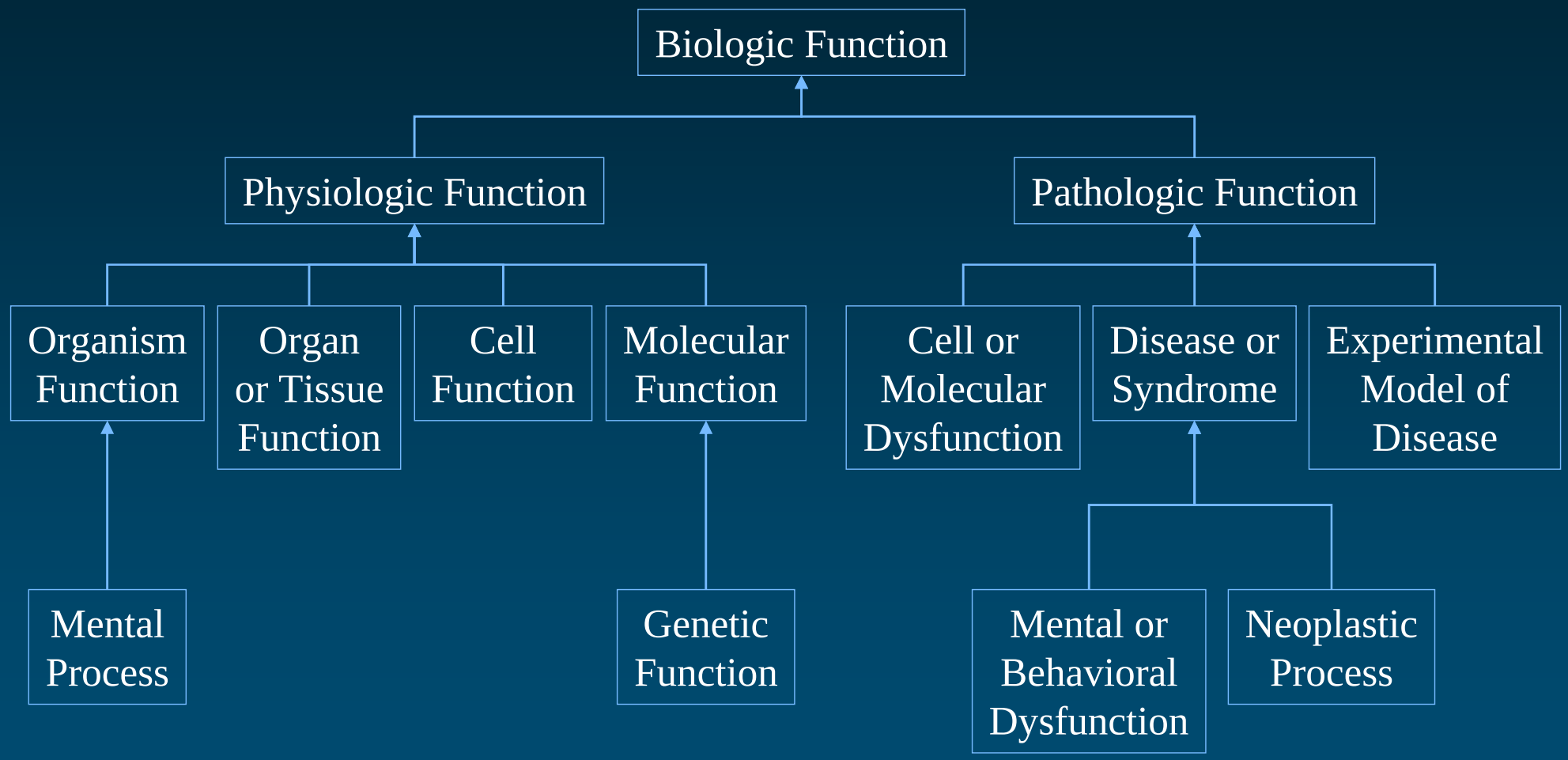
Semantic Network

- Semantic types (135)
  - tree structure
  - 2 major hierarchies
    - Entity
      - Physical Object
      - Conceptual Entity
    - Event
      - Activity
      - Phenomenon or Process

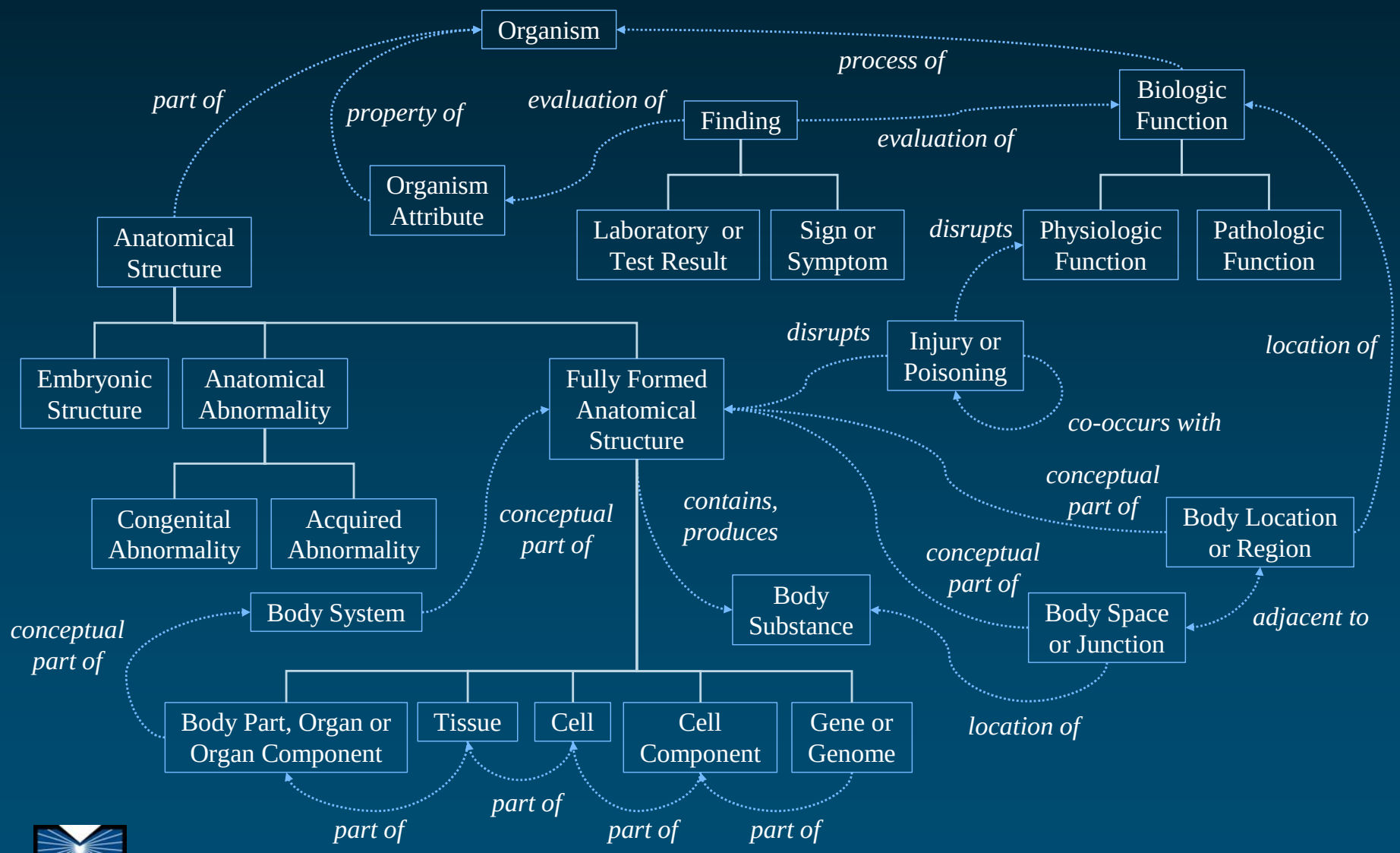
Semantic Network

- Semantic network relationships (54)
  - hierarchical (isa = is a kind of)
    - among types
      - Animal *isa* Organism
      - Enzyme *isa* Biologically Active Substance
    - among relations
      - treats *isa* affects
  - non-hierarchical
    - Sign or Symptom *diagnoses* Pathologic Function
    - Pharmacologic Substance *treats* Pathologic Function

“Biologic Function” hierarchy (isa)



Associative (non-isa) relationships

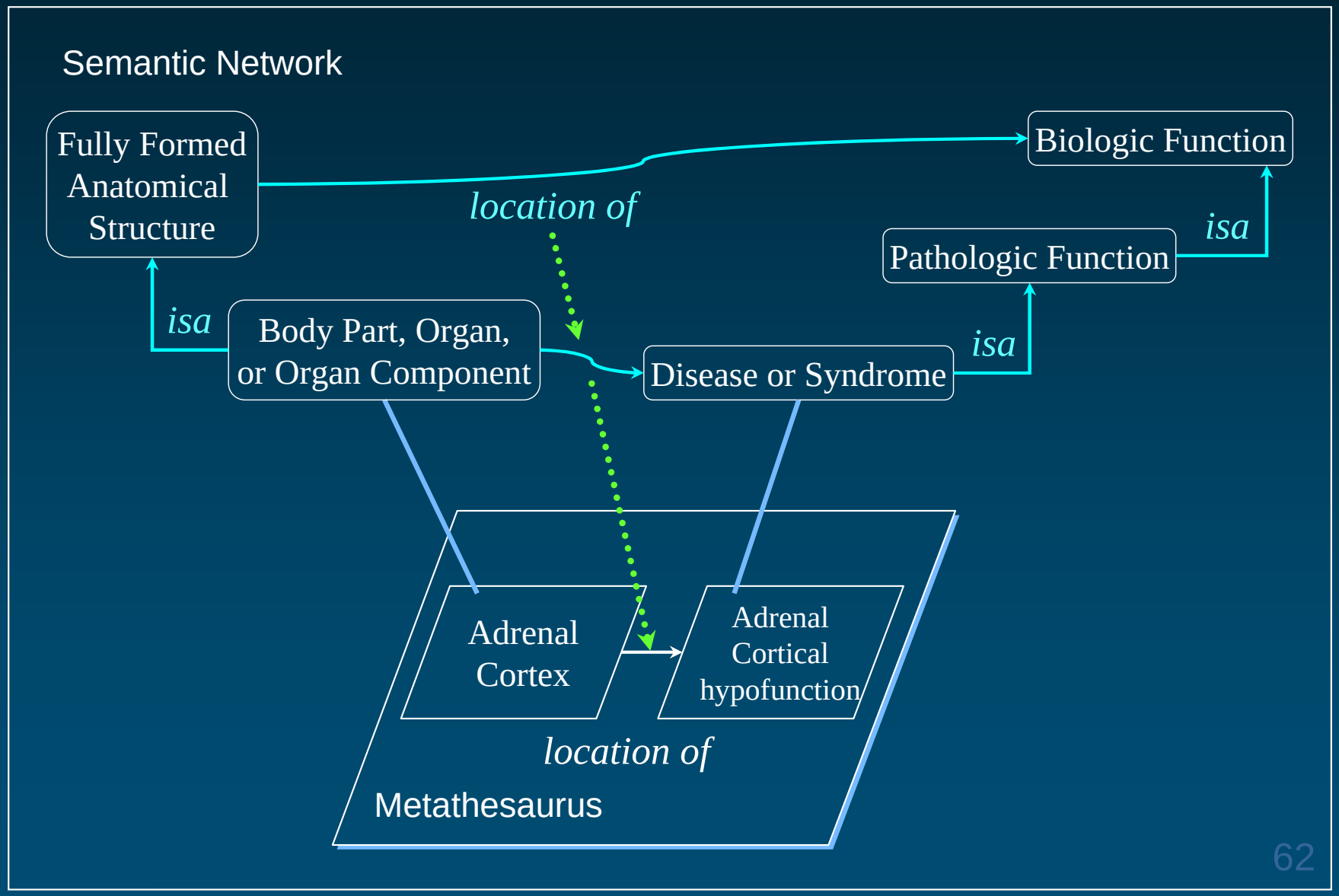


Why a semantic network?

- Semantic Types serve as high level categories assigned to Metathesaurus concepts, *independently of their position in a hierarchy*

- A relationship between 2 Semantic Types (ST) is a possible link between 2 concepts that have been assigned to those STs
  - The relationship may or may not hold at the concept level
  - Other relationships may apply at the concept level

Relationships can inherit semantics



SPECIALIST Lexicon and lexical tools

SPECIALIST Lexicon

- Content
  - English lexicon
  - Many words from the biomedical domain
- 360,000 lexical items
- Word properties
  - morphology
  - orthography
  - syntax
- Used by the lexical tools

Morphology

- Inflection
  - noun — nucleus, nuclei
  - verb — cauterize, cauterizes, cauterized, cauterizing
  - adjective — red, redder, reddest
- Derivation
  - verb ⟺ noun — cauterize -- cauterization
  - adjective ⟺ noun — red -- redness

Orthography

- Spelling variants
  - oe/e — oesophagus - esophagus
  - ae/e — anaemia - anemia
  - ise/ize — cauterise - cauterize
  - genitive mark — Addison's disease
Addison disease
Addisons disease

Syntax

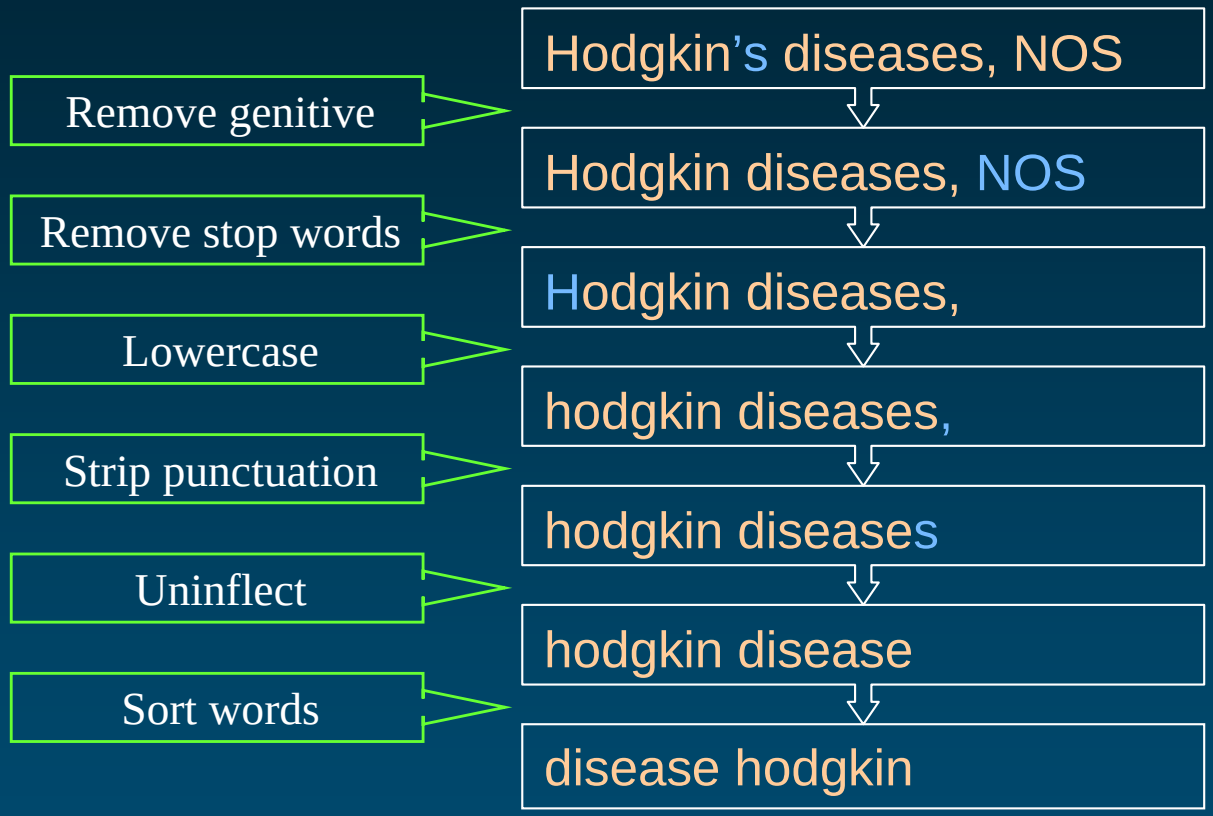
- Complementation
  - verbs
    - intransitive — I'll treat.
    - transitive — He treated the patient.
    - ditransitive — He treated the patient with a drug.
  - nouns
    - prepositional phrase

      Valve of coronary sinus
- Position for adjectives

Lexical tools

- To manage lexical variation in biomedical terminologies
- Major tools
  - Normalization
  - Indexes
  - Lexical Variant Generation program (lvg)
- Based on the SPECIALIST Lexicon
- Used by noun phrase extractors, search engines

Normalization



Normalization: Example

Hodgkin Disease
HODGKINS DISEASE
Hodgkin's Disease
Disease, Hodgkin's
Hodgkin's, disease
HODGKIN'S DISEASE
Hodgkin's disease
Hodgkins Disease
Hodgkin's disease NOS
Hodgkin's disease, NOS
Disease, Hodgkins
Diseases, Hodgkins
Hodgkins Diseases
Hodgkins disease
hodgkin's disease
Disease, Hodgkin

normalize → disease hodgkin

Normalization Applications

- Model for lexical resemblance
- Help find lexical variants for a term
  - Terms that normalize the same usually share the same LUI
- Help find candidates to synonymy among terms
- Help map input terms to UMLS concepts

Indexes

- Word index
  - word to Metathesaurus strings
  - one word index per language
- Normalized word index
  - normalized word to Metathesaurus strings
  - English only
- Normalized string index
  - normalized term to Metathesaurus strings
  - English only

Lexical Variant Generation program

- Tool for specialists (linguists)
- Performs atomic lexical transformations
  - generating inflectional variants
  - lowercase
  - …
- Performs sequences of atomic transformations
  - a specialized sequence of transformations provides the normalized form of a term (the *norm* program)

Part II

How to use the UMLS?

Outline

- Part II: *How to use the UMLS?*
  - Obtaining a license
  - Remote access
    - Knowledge Source Server (UMLSKS)
    - UMLSKS Application programming interface (API)
  - Local installation and customization (MetamorphoSys)
  - A UMLS-based algorithm: *Restrict to MeSH*
  - Benefits and limitations

Part II
How to use the UMLS?

(1) Obtaining a license

First step License agreement

- Online Web-based license:

  http://www.nlm.nih.gov/research/umls/license.html

  - Read license — Accept and continue
  - Read appendix 1 and 2 — Accept and continue
  - Print a copy for your records — Printer-friendly Version
  - Complete the Web form — Submit

  - Verify:
    - receive e-mail from NLM; go to Web site within 72 hours and enter first and last name
  - NLM official will countersign (turn-around time of a few days)
  - Receive 2nd e-mail from NLM with new license number

http://www.nlm.nih.gov/research/umls/license.html

United States
National Library of Medicine
National Institutes of Health

Search NLM Web Site
Go
NLM Home | Contact NLM | Site Map | FAQs

Unified Medical Language System

UMLS Home

Printer - friendly Version

This license contains the SNOMED CT® Affiliate Licence Agreement as Appendix 2. Users must accept this page, the contents of Appendix 1, and the contents of Appendix 2 to receive a UMLS license code.

License Agreement for Use of the UMLS® Metathesaurus®

This Agreement is made by and between the National Library of Medicine, Department of Health and Human Services (hereinafter referred to as "NLM") and the LICENSEE.

WHEREAS, the NLM was established by statute in order to assist the advancement of medical and related sciences, and to aid the dissemination and exchange of scientific and other information important to the progress of medicine and to the public health, (section 465 of the Public Health Service Act, as amended (42 U.S.C. section 286) and to carry out this purpose has been authorized to develop the Unified Medical Language System® (UMLS) to facilitate the retrieval and integration of machine-readable biomedical information from disparate sources;

WHEREAS, the NLM's UMLS project has produced the UMLS Metathesaurus, a machine-readable vocabulary knowledge source, that is useful in a variety of settings;

WHEREAS, the LICENSEE is willing to use the UMLS Metathesaurus at its sole risk and at no expense to NLM, which will result in information useful to NLM, may provide immediate improvements in biomedical information transfer to segments of the biomedical community, and is consistent with NLM's statutory functions,

NOW THEREFORE, it is mutually agreed as follows:

1. The NLM hereby grants a nonexclusive, non-transferable right to LICENSEE to use the UMLS Metathesaurus and incorporate its content in any computer applications or systems designed to improve access to biomedical information of any type subject to the restrictions in other provisions of this Agreement. The names and addresses of licensees auth[illegible] ducts are public information.
2. No charges, usage fees or royalties will be paid to NLM.

free from NLM

3. LICENSEE is prohibited from distributing the UMLS Metathesaurus or subsets of it, including individual vocabulary sources within the Metathesaurus, except (a) as an integral part of computer applications developed by LICENSEE for a purpose other than redistribution of vocabulary sources contained in the UMLS Metathesaurus and (b) if permitted by paragraph 12 of this agreement.
4. LICENSEE agrees to inform NLM prior to distributing any application(s) in which it is using the UMLS Metathesaurus and is encouraged to inform NLM of any difficulties encountered in using the UMLS Metathesaurus, and changes or enhancements to the UMLS Metathesaurus that would make it more useful to LICENSEE and its user groups.
5. Within 30 days of the end of any calendar year in which LICENSEE makes use of the UMLS Metathesaurus, LICENSEE agrees to provide NLM with a brief report on the usefulness of the UMLS Metathesaurus in general and, if applicable, on the usefulness of CPT i[illegible] format in particular. LICENSEE is strongly encouraged to submit to NLM locally developed extensions to the UMLS Metathe[illegible] potentially useful to other UMLS users for consideration for potential inclusion in the UMLS Metathesaurus.

annual report

6. NLM represents that the data provided under this Agreement were formatted with a reasonable standard of care, but makes no warranties express or implied, including no warranty of merchantability or fitness for particular purpose, regarding the accuracy or completeness of the data or that the machine-readable copy is error free. Therefore, LICENSEE agrees to hold NLM, the Government, and any organization contributing a vocabulary source to the UMLS Metathesaurus free from any liability resulting from errors in terminology or other data or on the machine-readable copy. NLM and such other organizations disclaim any liability for any consequences due to use, misuse, or interpretation of information contained or not contained in the UMLS Metathesaurus.
7. NLM represents that its ability to continue to include certain vocabulary sources within the UMLS Metathesaurus is dependent on continuing contractual relations or agreements with the copyright holders for these vocabulary sources. Therefore, LICENSEE agrees to hold NLM and the individual copyright holder free from any liability resulting from the removal of any vocabulary source from future editions of the UMLS Metathesaurus.
8. NLM reserves the right to change the type and format of its machine-readable data. NLM agrees to inform LICENSEE of any changes to the format of the UMLS Metathesaurus, EXCEPT the addition of entirely new data elements to the Metathesaurus, at least 90 days before the data are distributed.
9. The presence in the UMLS Metathesaurus of vocabulary or data produced by organizations other than NLM does not imply any endorsement of the UMLS Metathesaurus by these organizations.
10. LICENSEE shall acknowledge NLM as its source of the UMLS Metathesaurus, citing the year and version number, in a suitable and customary manner but may not in any way indicate or imply that NLM or any of the organizations whose vocabulary sources are included in the UMLS has endorsed LICENSEE or its products.

11. Some of the Material in the UMLS Metathesaurus is from copyrighted sources. If LICENSEE uses any material from copyrighted sources from the UMLS Metathesaurus:

a) the LICENSEE is required to display in full, prior to providing user access to the Metathesaurus or any of the vocabulary sources within the UMLS, the following wording in order that its users be made aware of these copyright constraints:

> "Some material in the UMLS Metathesaurus is from copyrighted sources of the respective copyright holders. Users of the UMLS Metathesaurus are solely responsible for compliance with any copyright, patent or trademark restrictions and are referred to the copyright, patent or trademark notices appearing in the original sources, all of which are hereby incorporated by reference.";

to display a list of all of the vocabularies contained within the UMLS Metathesaurus that are used in the LICENSEE's application; and to indicate for each vocabulary any appropriate copyright notice and whether the entire contents is present or only a portion of it.

b) the LICENSEE is prohibited from altering UMLS and other vocabulary source content contained within the UMLS Metathesaurus, but may include content from other sources in applications that also contain content from the UMLS Metathesaurus. The LICENSEE may not imply in any way that data from other sources is part of the UMLS Metathesaurus or of any

need to retain UMLS identifiers

c) the LICENSEE is required to include in its applications identifiers from the UMLS Metathesaurus such that the original source vocabularies for any data obtained from the UMLS Metathesaurus can be determined by reference to a complete version of the UMLS Metathesaurus.

12. For material in the UMLS Metathesaurus obtained from some sources additional restrictions on LICENSEE's use may apply. The categories of additional restrictions are described below. The list of UMLS Metathesaurus Vocabulary Sources, which is part of this Agreement and is updated when each version of the Metathesaurus is released, indicates the category of additional restrictions, if any, that apply to each vocabulary source.

LICENSEE should contact the copyright holder directly to discuss uses of a source vocabulary beyond those allowed under this license agreement. If LICENSEE or LICENSEE's end user has a separate agreement with the copyright holder for use of a UMLS Metathesaurus source vocabulary, LICENSEE or LICENSEE's end user may use vocabulary source content obtained from the UMLS Metathesaurus in accordance with the terms of the separate agreement.

additional licenses may be necessary

12. 1. **Category 1:**

LICENSEE is prohibited from translating the vocabulary source into another language or from producing other derivative works based on this single vocabulary source.

12. 2. **Category 2:**

All category 1 restrictions AND

LICENSEE is prohibited from using the vocabulary source in operational applications that create records or information containing data from the vocabulary source. Use for data creation research or product development is allowed.

12. 3. **Category 3:**

LICENSEE's right to use material from the source vocabulary is restricted to internal use at the LICENSEE's site(s) for research, product development, and statistical analysis only. Internal use includes use by employees, faculty, and students of a single institution at multiple sites. Notwithstanding the foregoing, use by students is limited to doing research under the direct supervision of faculty. Internal research, product development, and statistical analysis use expressly excludes: use of material from these copyrighted sources in routine patient data creation; incorporation of material from these copyrighted sources in any publicly accessible computer-based information system or public electronic bulletin board including the Internet; publishing or translating or creating derivative works from material from these copyrighted sources; selling, leasing, licensing, or otherwise making available material from these copyrighted works to any unauthorized party; and copying for any purpose except for back up or archival purposes.

LICENSEE may be required to display special copyright, patent and/or trademark notices before displaying content from the vocabulary source. Applicable notices are included in the list of UMLS Metathesaurus Vocabulary sources, that is part of this Agreement.

12. 4. **Category 4:**

special category for SNOMED CT

12. 4. **Category 4:**

12.4.1. LICENSEE is prohibited from translating the vocabulary source into another language or from altering the vocabulary source content.

12.4.2. LICENSEE's right to use the vocabulary source is restricted to use in the U.S. by LICENSEE's employees, contractors, faculty, students, clients, patients, or constituents within electronic systems or devices built, purchased, licensed, or used by LICENSEE for U.S. governmental purposes or for any health care, public health, research, educational, or statistical use in the U.S. Use by students is limited to research or educational activities under the direct supervision of faculty.

12.4.3. LICENSEE has the right to distribute the vocabulary source in the U.S., but only in combination with other UMLS Metathesaurus content. Further, LICENSEE's right to distribute is restricted to:

a. Electronic distribution to LICENSEE's direct U.S. affiliates, or to other U.S. entities that have signed the UMLS license, in order to facilitate use of the vocabulary for health care, public health, research, educational or statistical purposes in the U.S. only.

    i. LICENSEE must take reasonable precautions to prevent distribution of the vocabulary source to non-US entities.

    ii. LICENSEE must include in its annual report a list of all U.S. affiliates or other U.S. entities to whom it has distributed content from the vocabulary source.

b. Distribution of encoded patient level data sets or knowledge encoded in the vocabulary source by LICENSEE to any U.S. entity for use in the U.S. only.
c. Inclusion of encoded records or content from the vocabulary source in: (1) free publicly accessible retrieval systems or (2) fee-based retrieval systems that are accessible within the U.S. only, provided that these systems do not permit users to copy or extract any significant portion of the vocabulary source.

12.4.4. DEFINITIONS

a. U.S. is defined as all U.S. states, territories, and the District of Columbia; any U.S. government facility or office, whether permanent or temporary, wherever located; and access to a system in any of these locations by U.S. government employees, designated representatives or contractors, wherever located, for U.S. government purposes.

b. U.S. entity is defined as (i) for government entities, an agency or department of the U.S. Government, (ii) for corporations, as a corporation incorporated and operating in the U.S.; and (iii) for other entities as an entity organized under the laws of the U.S.

13. LICENSEE shall take reasonable steps to ensure that anyone who has authorized access to data or vocabulary sources from the UMLS Metathesaurus under this Agreement complies with its provisions.

14. LICENSEE and/or its end users shall be solely responsible for compliance with any copyright or other restrictions on vocabulary sources in the UMLS Metathesaurus; NLM assumes no responsibility or liability associated with the LICENSEE's (or any of the LICENSEE's users) use and/or reproduction of copyrighted material, patent or trademark violations. Anyone contemplating reproduction of all or any portion of the UMLS Metathesaurus or any of its vocabulary sources should consult legal counsel.

    The holder of a copyright in any vocabulary source shall be a third party beneficiary to this agreement and shall have a right to enforce the agreement against any LICENSEE that violates any provision pertaining to that copyright holder.

15. This Agreement shall be effective until terminated by one of the parties upon 30 days written notice to the other party. LICENSEE's failure to abide by the terms of the Agreement shall be grounds for its termination. Neither the Government, its employees, or any vocabulary sources contained in the UMLS Metathesaurus shall be liable or responsible to LICENSEE in any manner whatsoever for damages of any nature whatsoever arising from the termination of this Agreement.

16. In the event that any provision of this Agreement is determined to violate any law or is unenforceable, the remainder of the Agreement shall remain in full force and effect.

Accept and continue | Not accept

Accept & continue

Printer - friendly Version

APPENDIX A.1

Appendix to the License Agreement for Use of the UMLS® Metathesaurus

UMLS METATHESAURUS® SOURCE VOCABULARIES - 2007AB Edition

Sources are listed in order according to the abbreviations used in the UMLS Metathesaurus files. If additional restrictions and notices apply, the category of restrictions and the special notices appear under the name of the source. See the license agreement for an explanation of the categories of restrictions. Many sources publish printed editions and/or other explanatory information that may be essential to understanding the purpose and application of particular sources in data creation and retrieval. Contact information is provided for each source. Please address questions about permissions or license agreements for additional uses not covered by this Agreement, or other inquiries about individual sources, to the appropriate contacts.

NLM is working toward inclusion in the UMLS Metathesaurus of the complete, current edition of most of these vocabulary sources.

AIR93 - AI/RHEUM. Bethesda, (MD) : National Library of Medicine, Lister Hill Center, 1993.

Contact: May Cheh, Lister Hill National Center for Biomedical Communications, National Library of Medicine, Building 38A, Room 9E902, 8600 Rockville Pike, Bethesda, MD 20894; e-mail: cheh@nlm.nih.gov.

ALT2006 - Alternative Billing Concepts (Altlink). 7th Version. Las Cruces, NM, 2006.

CATEGORY 3 RESTRICTIONS APPLY

source restriction level

Contact: Alternative Link LLC, 6121 Indian School Road NE, Suite 131, Albuquerque, NM 87110; Phone: (505) 875-0001; Toll Free: (877)621-5465; Fax: (505) 875-0002; e-mail: ail@alternativelink.com

---

• • •

• • •

WHOFRE_1997 - WHO Adverse Drug Reaction Terminology (WHOART). French Translation. Uppsala (Sweden): WHO Collaborating Centre for International Drug Monitoring, 1997.

CATEGORY 2 RESTRICTIONS APPLY

Contact: WHO Collaborating Centre for International Drug Monitoring, Stora Target 3, S-753 20 Uppsala, Sweden; fax: +46-18-656080

WHOGER_1997 - WHO Adverse Drug Reaction Terminology (WHOART). German Translation. Uppsala (Sweden): WHO Collaborating Centre for International Drug Monitoring, 1997.

CATEGORY 2 RESTRICTIONS APPLY

Contact: WHO Collaborating Centre for International Drug Monitoring, Stora Target 3, S-753 20 Uppsala, Sweden; fax: +46-18-656080

WHOPOR_1997 - WHO Adverse Drug Reaction Terminology (WHOART). Portuguese Translation. Uppsala (Sweden): WHO Collaborating Centre for International Drug Monitoring, 1997.

CATEGORY 2 RESTRICTIONS APPLY

Contact: WHO Collaborating Centre for International Drug Monitoring, Stora Target 3, S-753 20 Uppsala, Sweden; fax: +46-18-656080

WHOSPA_1997 - WHO Adverse Drug Reaction Terminology (WHOART). Spanish Translation. Uppsala (Sweden): WHO Collaborating Centre for International Drug Monitoring, 1997.

CATEGORY 2 RESTRICTIONS APPLY

Contact: WHO Collaborating Centre for International Drug Monitoring, Stora Target 3, S-753 20 Uppsala, Sweden; fax: +46-18-656080

Accept and continue | Not accept

Printer - friendly Version

APPENDIX 2

SNOMED CT® AFFILIATE LICENCE AGREEMENT

IMPORTANT NOTICE – PLEASE READ THE FOLLOWING CAREFULLY

This is a Licence Agreement between (1) **The International Health Terminology Standards Development Organisation** (forening med begrænset ansvar), an association (foreningen) established under the laws of Denmark, whose principal place of business is at Rued Langgaards Vej 7, 5te A22, DK-2300 Copenhagen S, Denmark (the "**Licensor**") and (2) the person or organisation to which the International Release of SNOMED CT (whether on its own or as part of a Member's National Release of SNOMED CT) is distributed or otherwise made available (the "**Licensee**").

By downloading, accessing or using any part of the International Release of SNOMED CT or a Member's National Release of SNOMED CT, or exercising any rights granted under this Licence Agreement, the Licensee agrees to be bound by the terms of this Licence Agreement.

• • •

License Restriction Levels 0-4 (2007AB)

- Level 0 (79.3%)
  - *unrestricted*
- Level 1 (5.0%)
  - *negotiate to translate*
- Level 2 (0.5%)
  - *negotiate to use in health data creation*
- Level 3 (25.4%)
  - *negotiate to use in production*
  - *explicitly prohibited to provide Internet access*
- Level 4 (21.8%)
  - *SNOMED CT (unrestricted in member countries)*

There may be additional restrictions, or separate license fees, associated with usage of specific vocabularies. Read the UMLS License, including the Appendix!

NLM

Part II
How to use the UMLS?

(2) Remote access

Remote Access

- UMLS Knowledge Source Server:

  http://umlsks.nlm.nih.gov

- Web search interface
- Application Programming Interface (API)

- Coming soon: web services

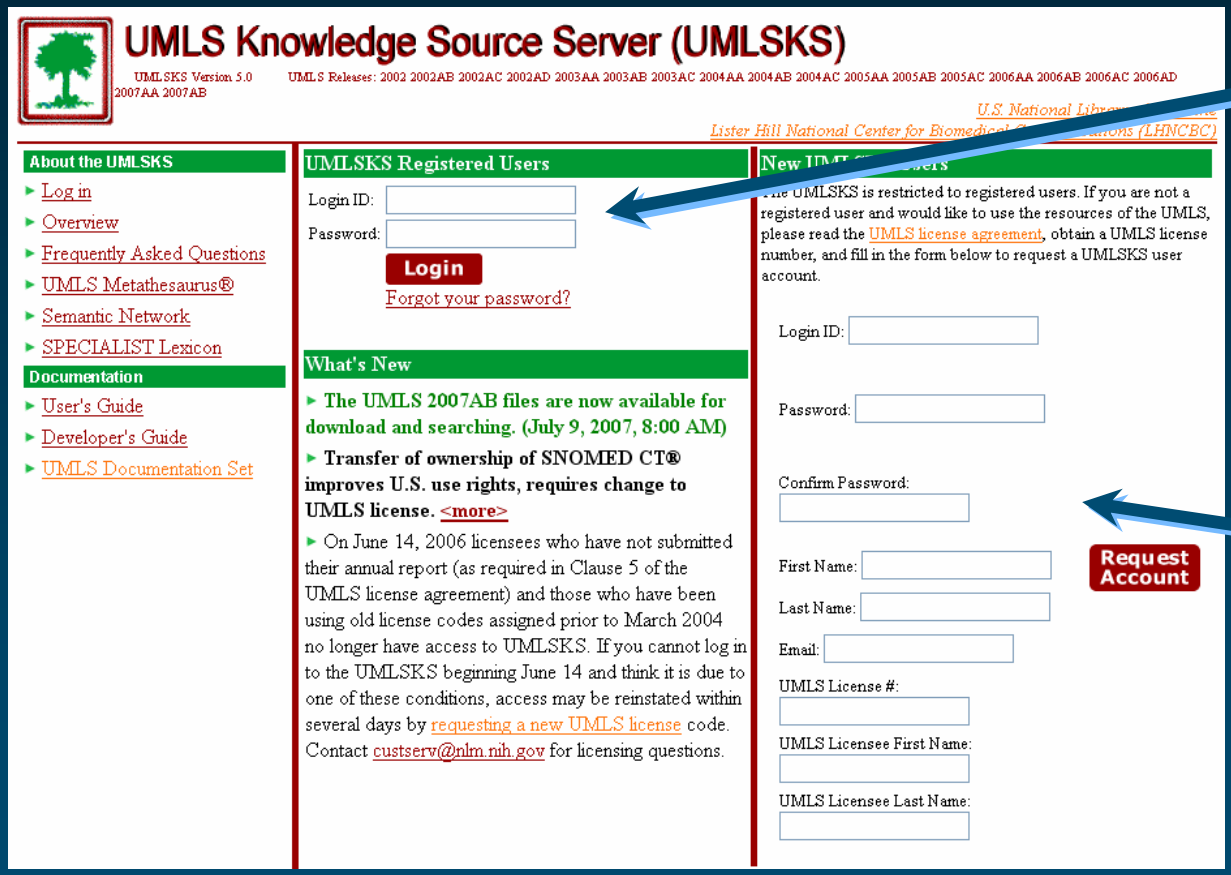
Knowledge Source Server

Web search interface

UMLSKS Web search interface

- Logging in
- Basic searching
- Advanced searching

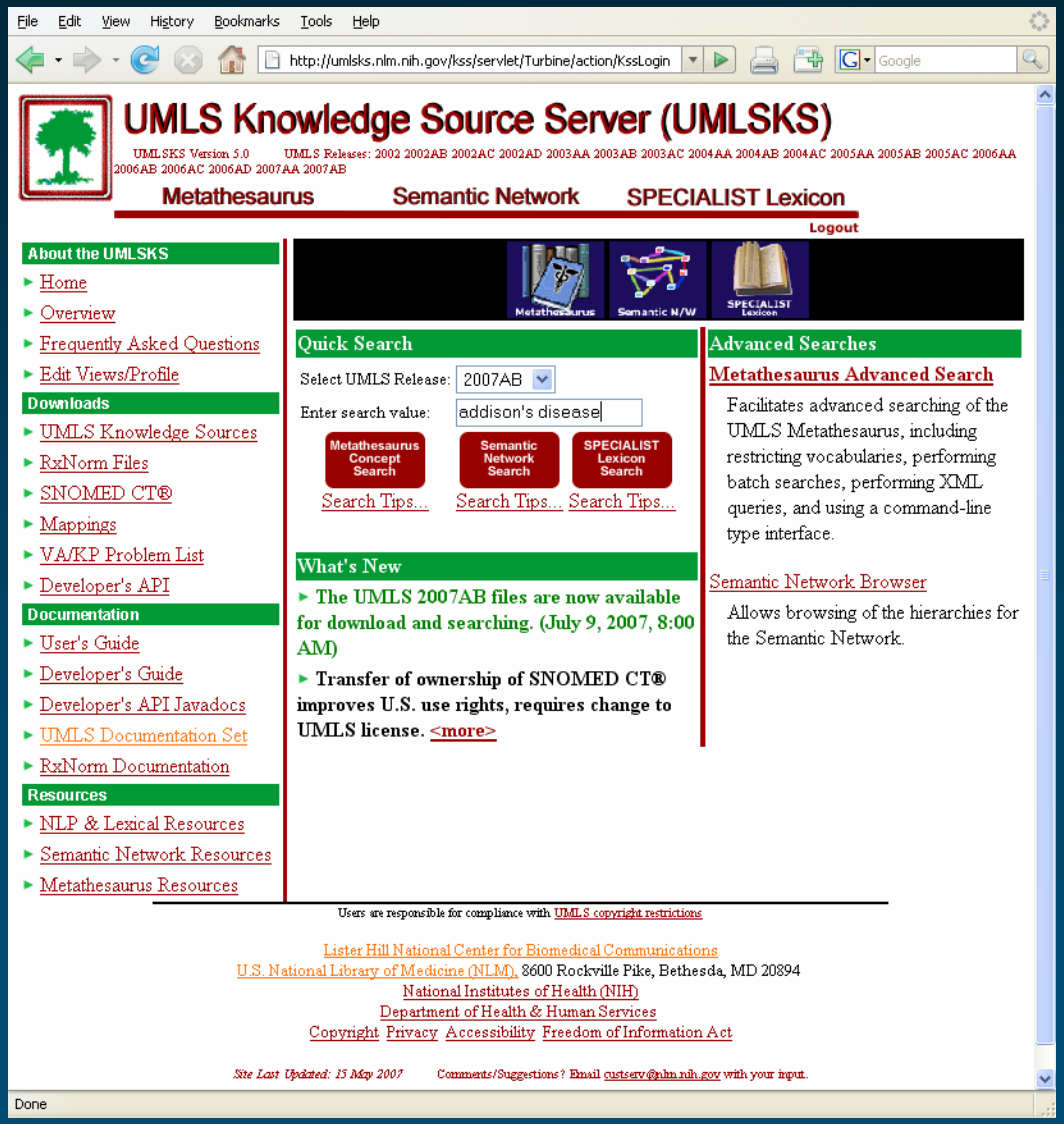
UMLSKS Web search interface log in



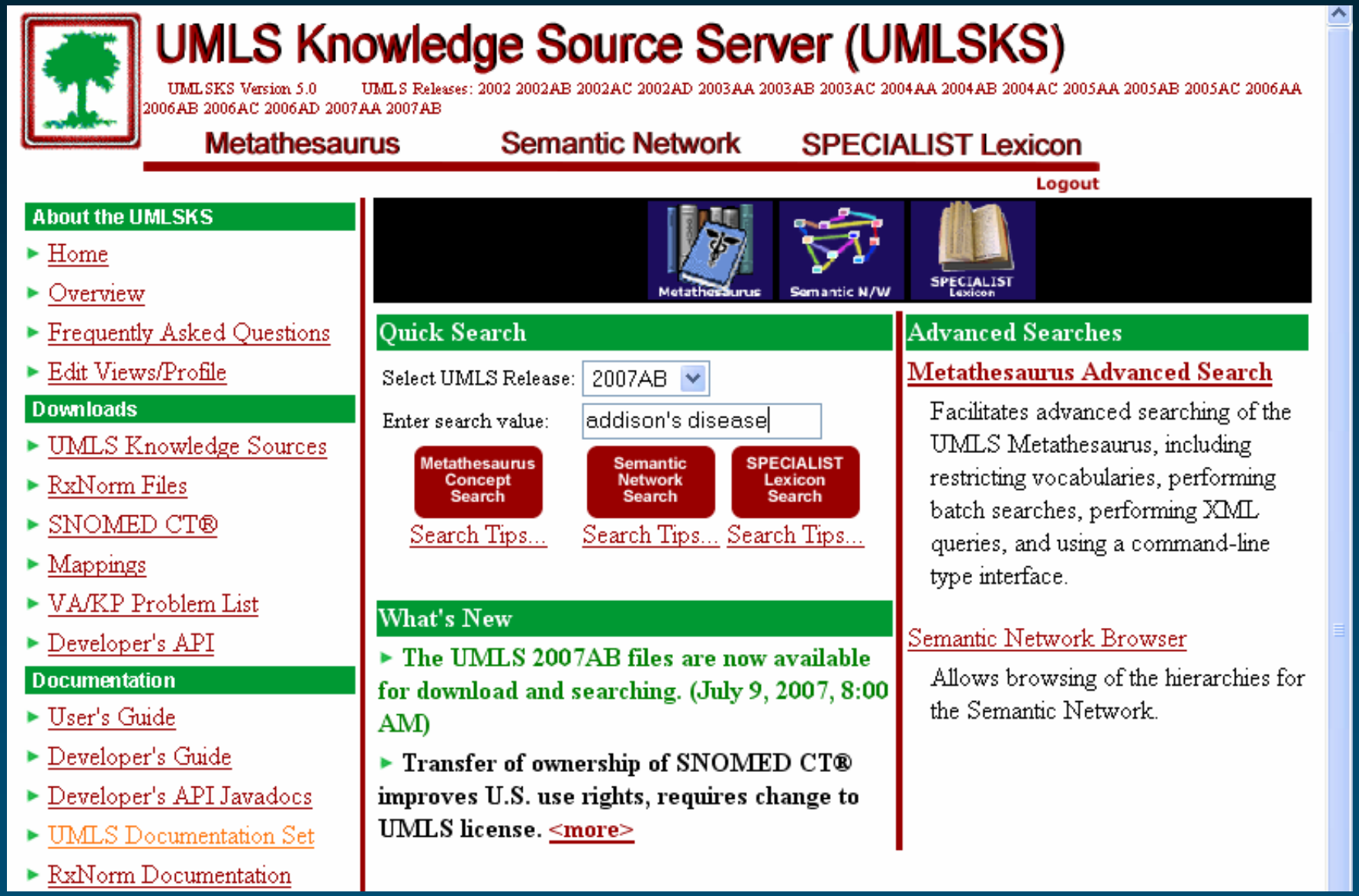
- Returning users log in
- New users create account

UMLS Knowledge Source Server Home Page

- Tabs across top access basic searching of 3 Knowledge Sources
- Advanced searching options on right-hand side

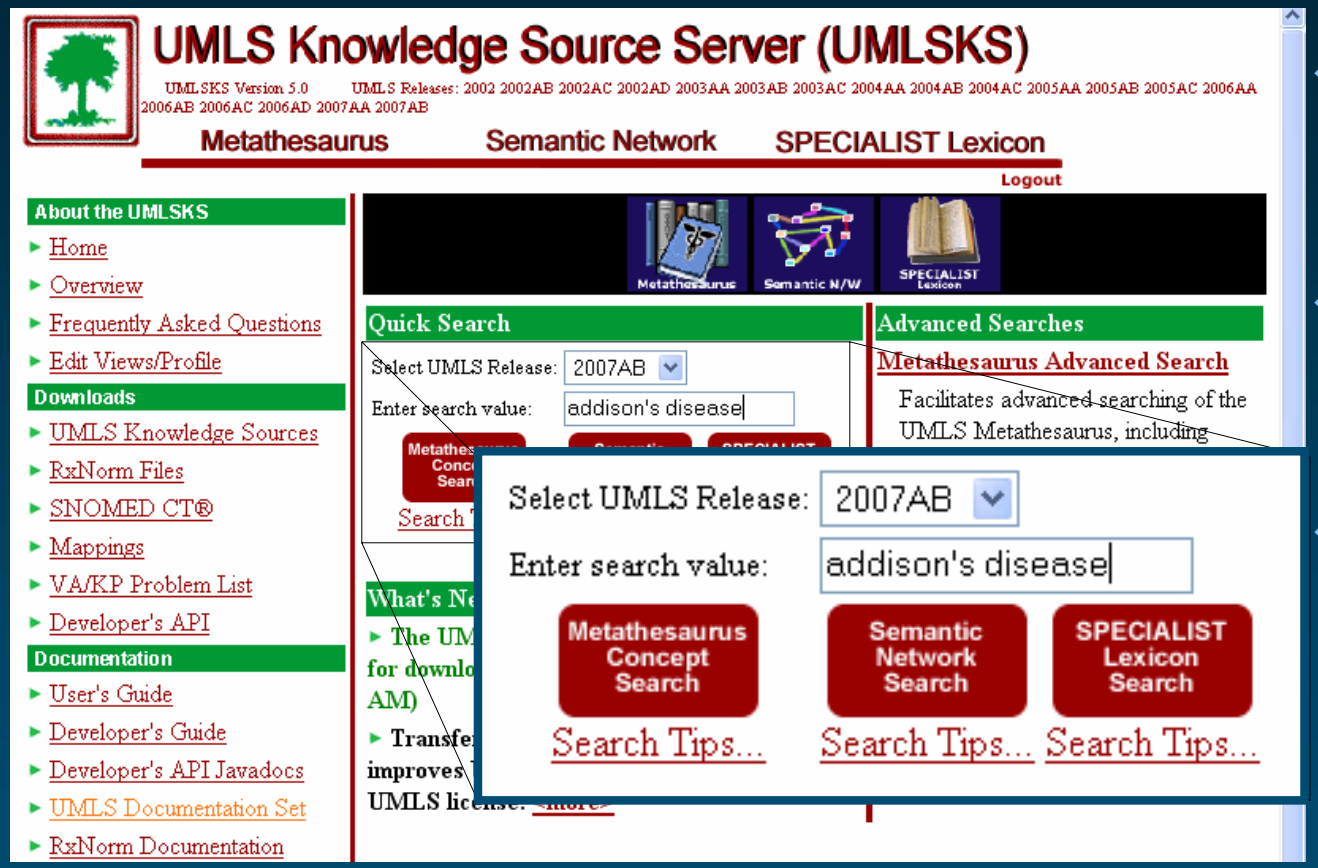


UMLS Knowledge Source Server Home Page



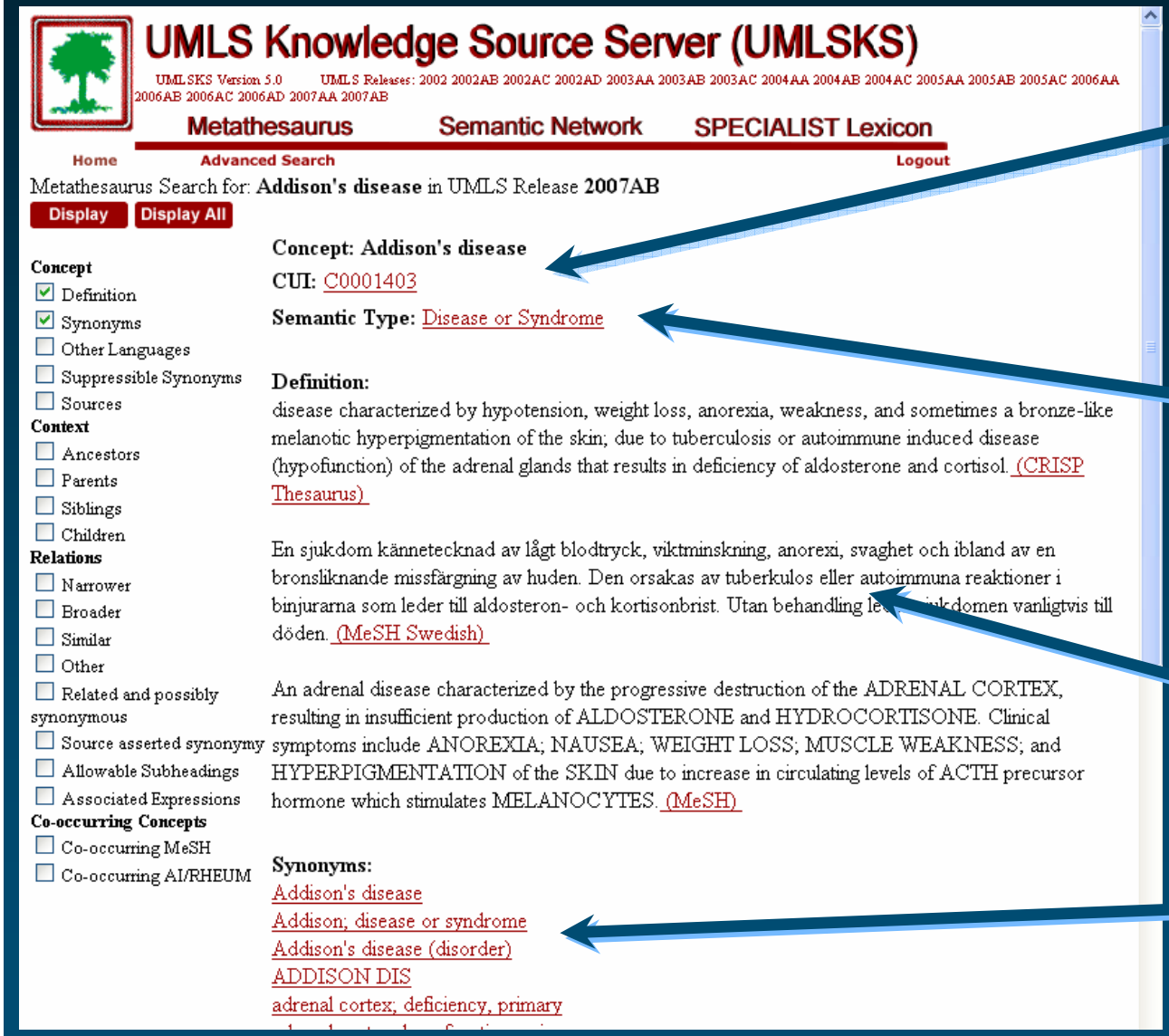
Metathesaurus Basic Search

Addison's disease



- UMLS Release
- Search Term
- UMLS Knowledge Source

Concept Report *Addison's disease*



UMLS Knowledge Source Server (UMLSKS)

UMLSKS Version 5.0 UMLS Releases: 2002 2002AB 2002AC 2002AD 2003AA 2003AB 2003AC 2004AA 2004AB 2004AC 2005AA 2005AB 2005AC 2006AA 2006AB 2006AC 2006AD 2007AA 2007AB

Metathesaurus Semantic Network SPECIALIST Lexicon

Home Advanced Search Logout

Metathesaurus Search for: **Addison's disease** in UMLS Release **2007AB**

Display Display All

Concept
- Definition
- Synonyms
- Other Languages
- Suppressible Synonyms
- Sources

Context
- Ancestors
- Parents
- Siblings
- Children

Relations
- Narrower
- Broader
- Similar
- Other
- Related and possibly synonymous
- Source asserted synonymy
- Allowable Subheadings
- Associated Expressions

Co-occurring Concepts
- Co-occurring MeSH
- Co-occurring AI/RHEUM

Concept: Addison's disease

CUI: C0001403

Semantic Type: Disease or Syndrome

Definition:

disease characterized by hypotension, weight loss, anorexia, weakness, and sometimes a bronze-like melanotic hyperpigmentation of the skin; due to tuberculosis or autoimmune induced disease (hypofunction) of the adrenal glands that results in deficiency of aldosterone and cortisol. (CRISP Thesaurus)

En sjukdom kännetecknad av lågt blodtryck, viktminskning, anorexi, svaghet och ibland av en bronsliknande missfärgning av huden. Den orsakas av tuberkulos eller autoimmuna reaktioner i binjurarna som leder till aldosteron- och kortisonbrist. Utan behandling le[illegible]jukdomen vanligtvis till döden. (MeSH Swedish)

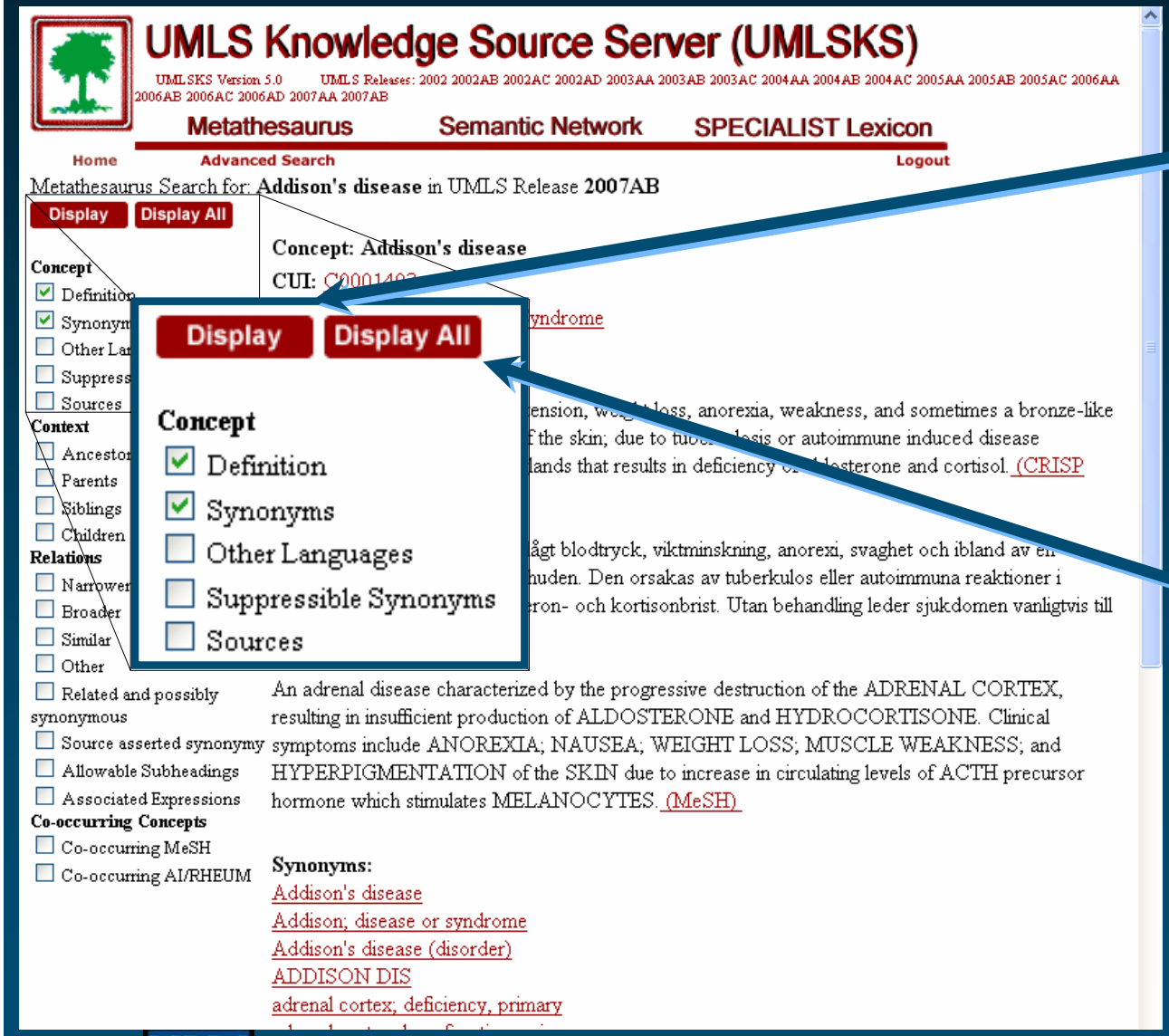
An adrenal disease characterized by the progressive destruction of the ADRENAL CORTEX, resulting in insufficient production of ALDOSTERONE and HYDROCORTISONE. Clinical symptoms include ANOREXIA; NAUSEA; WEIGHT LOSS; MUSCLE WEAKNESS; and HYPERPIGMENTATION of the SKIN due to increase in circulating levels of ACTH precursor hormone which stimulates MELANOCYTES. (MeSH)

Synonyms:

Addison's disease
Addison; disease or syndrome
Addison's disease (disorder)
ADDISON DIS
adrenal cortex; deficiency, primary

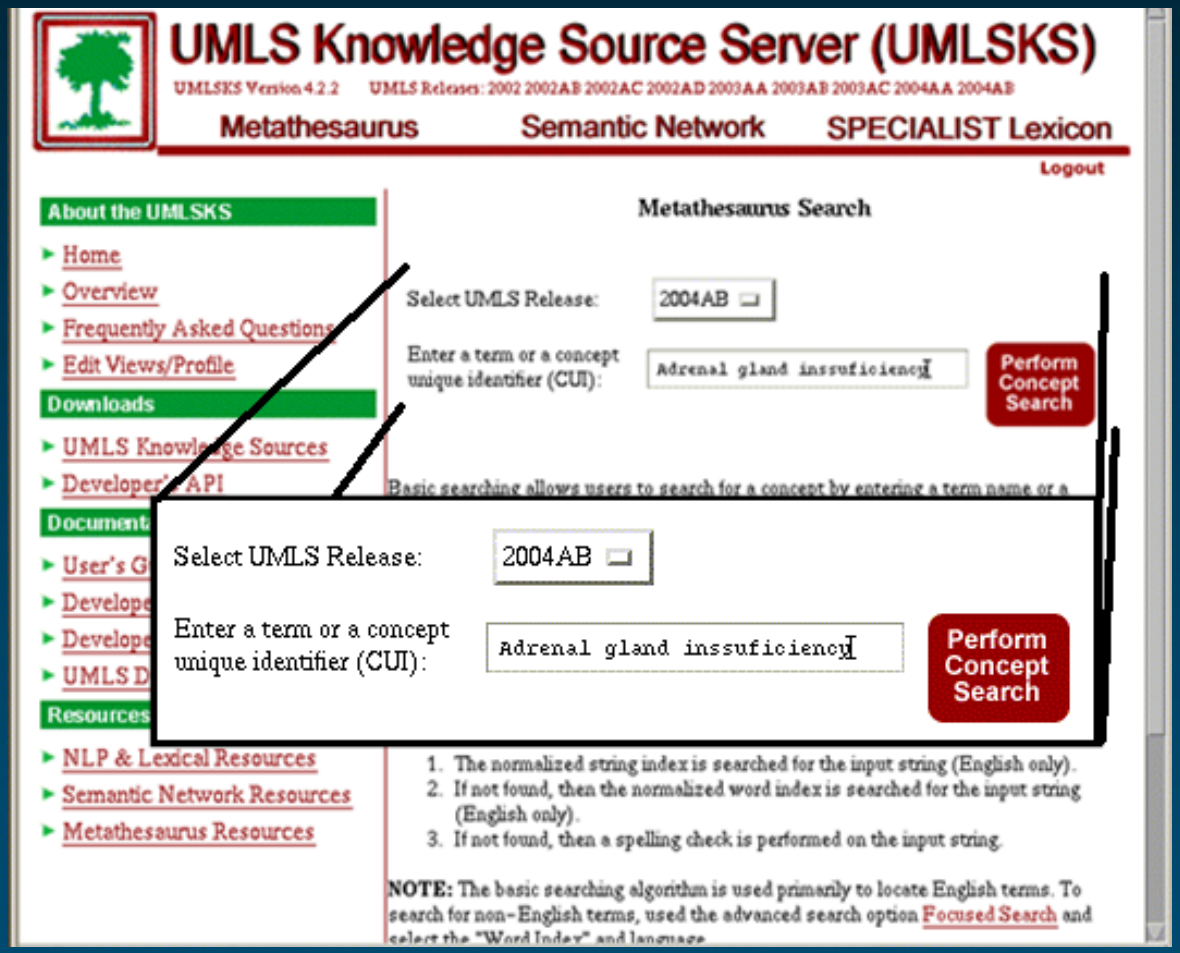
- Concept Name /CUI
- Semantic Type(s)
- Definition(s)
- Synonyms

Display All



- "Display" shows results for selected options
- "Display All" shows results for all available options

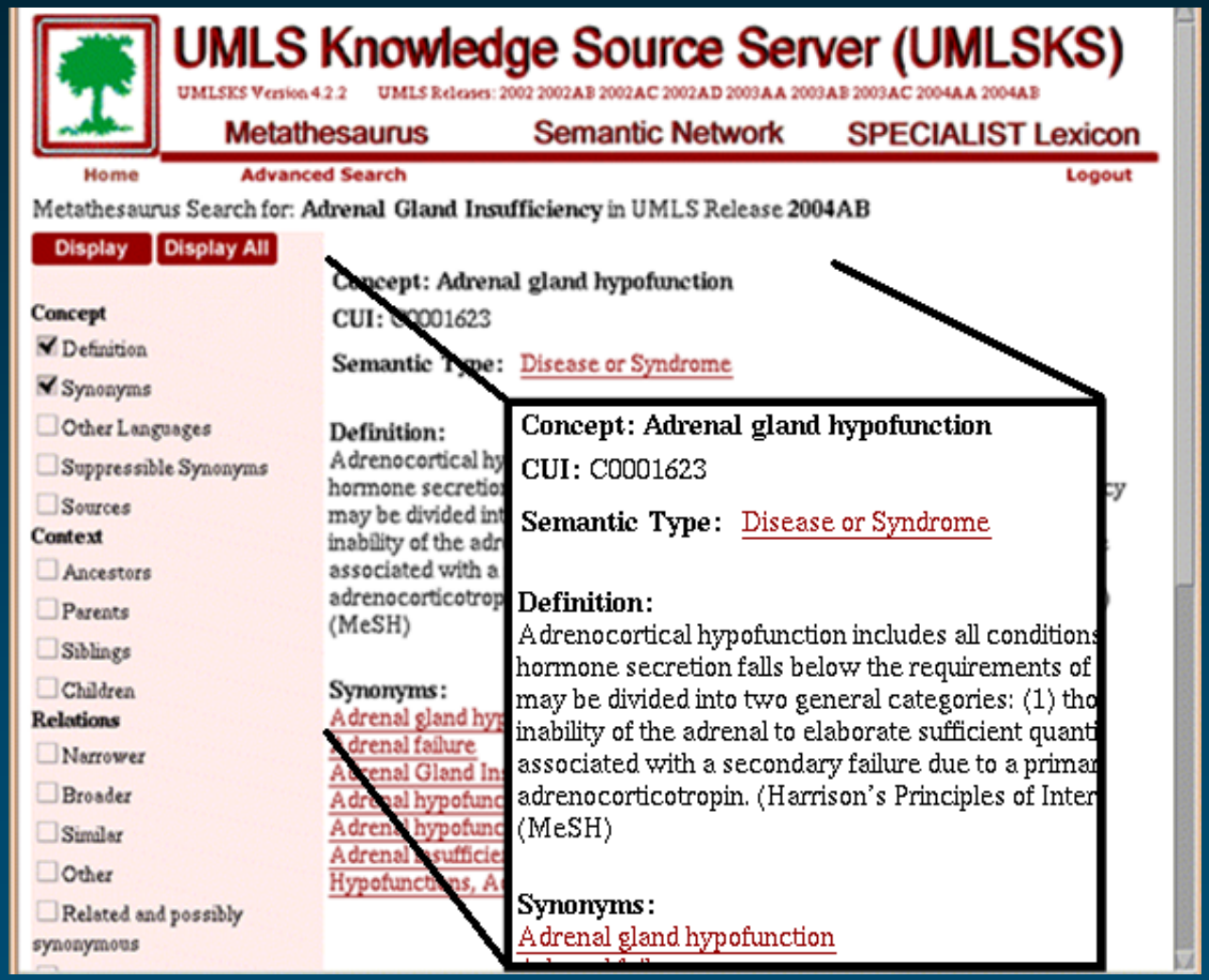
Metathesaurus Basic Search
Adrenal gland insufficiency



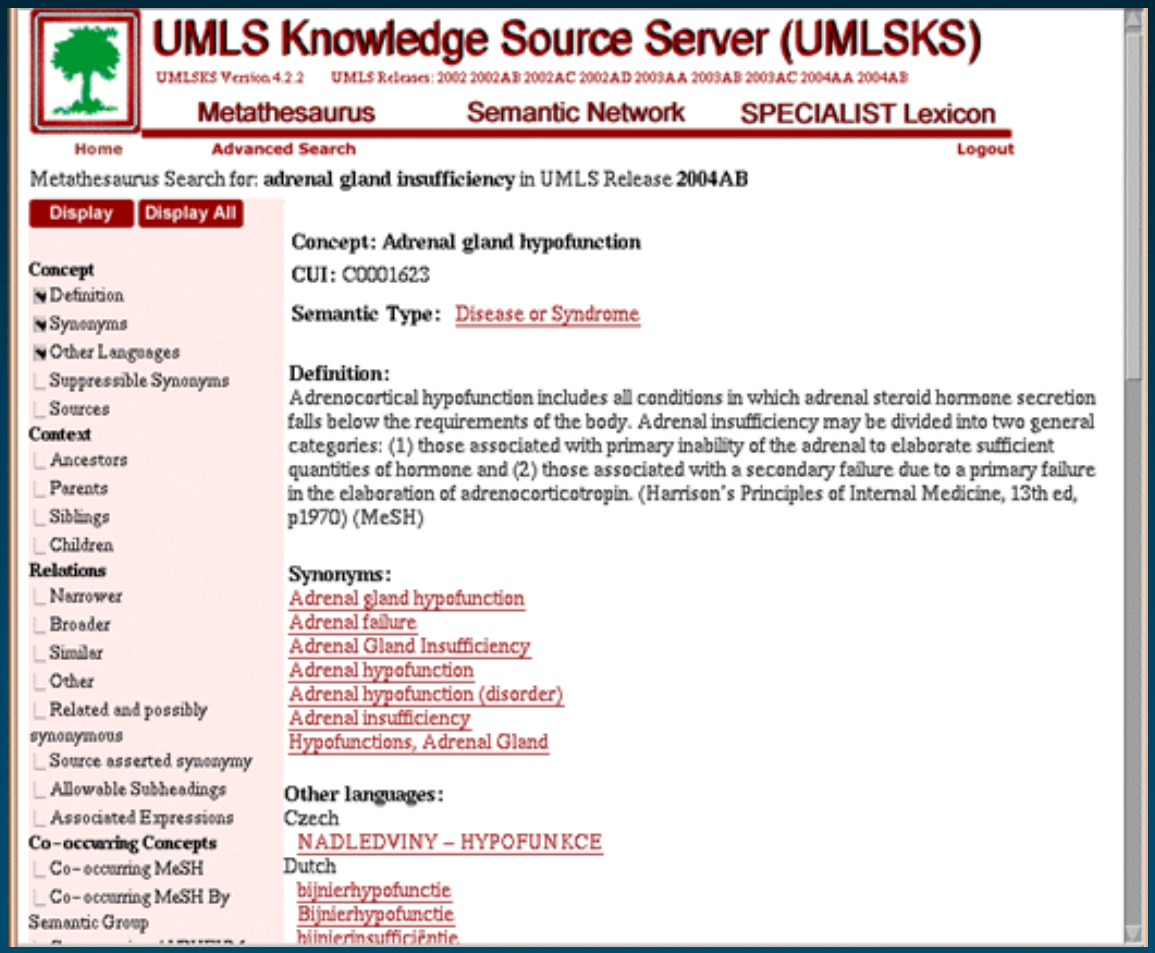
- Specify:
  - UMLS Release
  - Search term
- Algorithm:
  - Search Normalized String
  - Search Normalized Word
  - Suggest Spelling

Basic Concept Report

Adrenal gland insufficiency



Concept Report Display All
Adrenal Gland Insufficiency

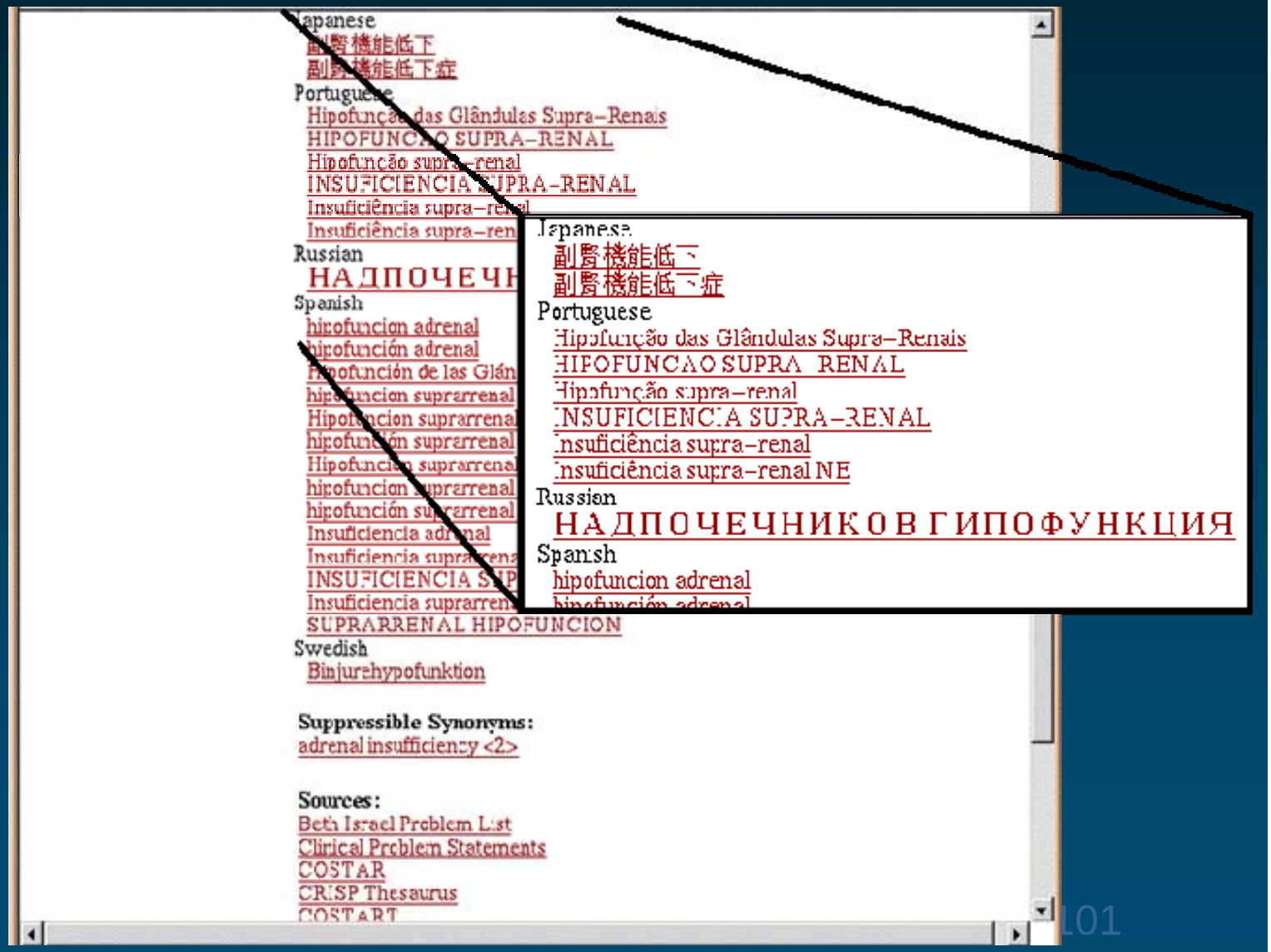


- Concept Name/CUI
- Semantic Type(s)
- Definition(s)
- Synonyms, including foreign languages
- Relations (broader, narrower, etc.)
- Co-occurrence data

Concept Report Display All (continued)

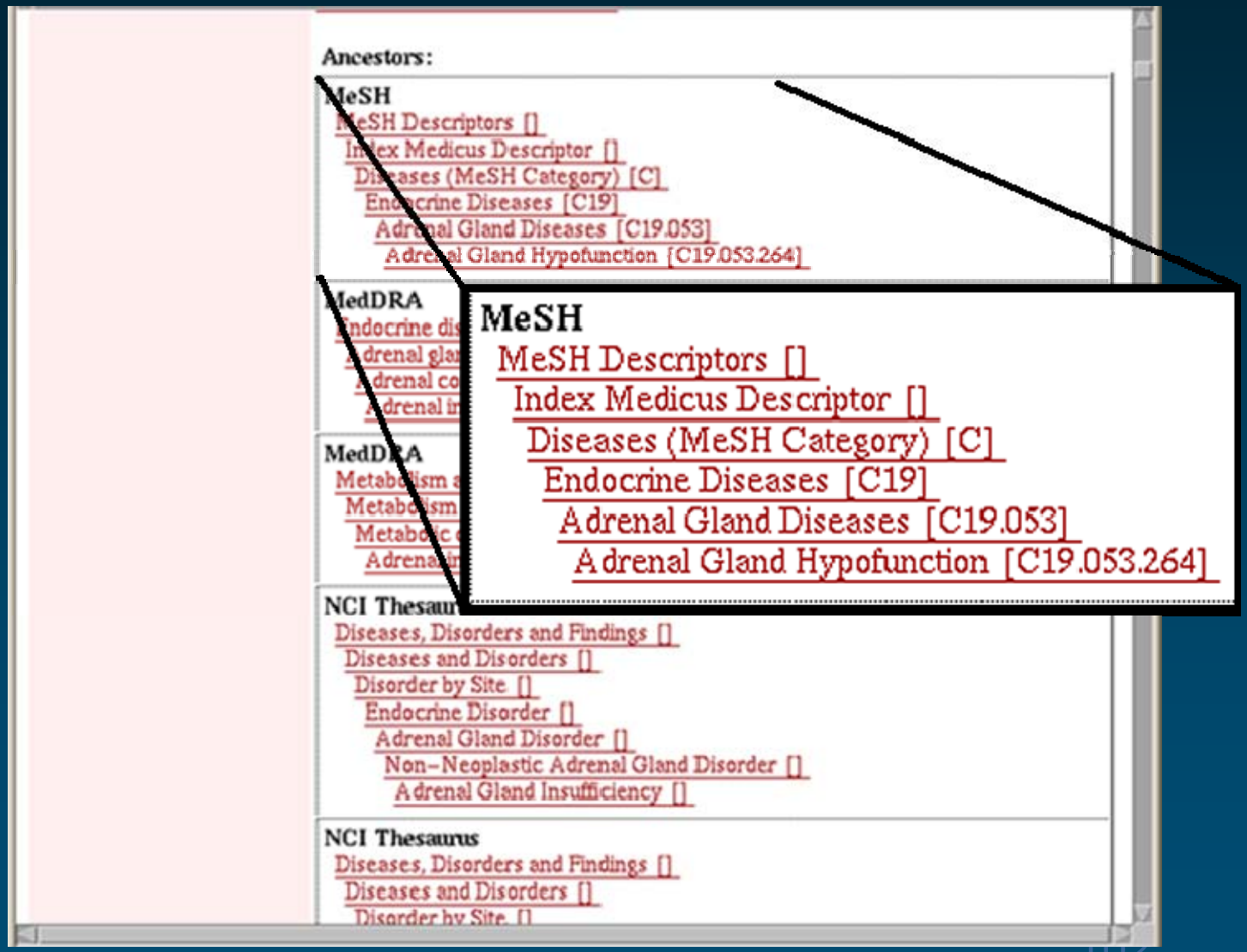
Synonyms

Sources



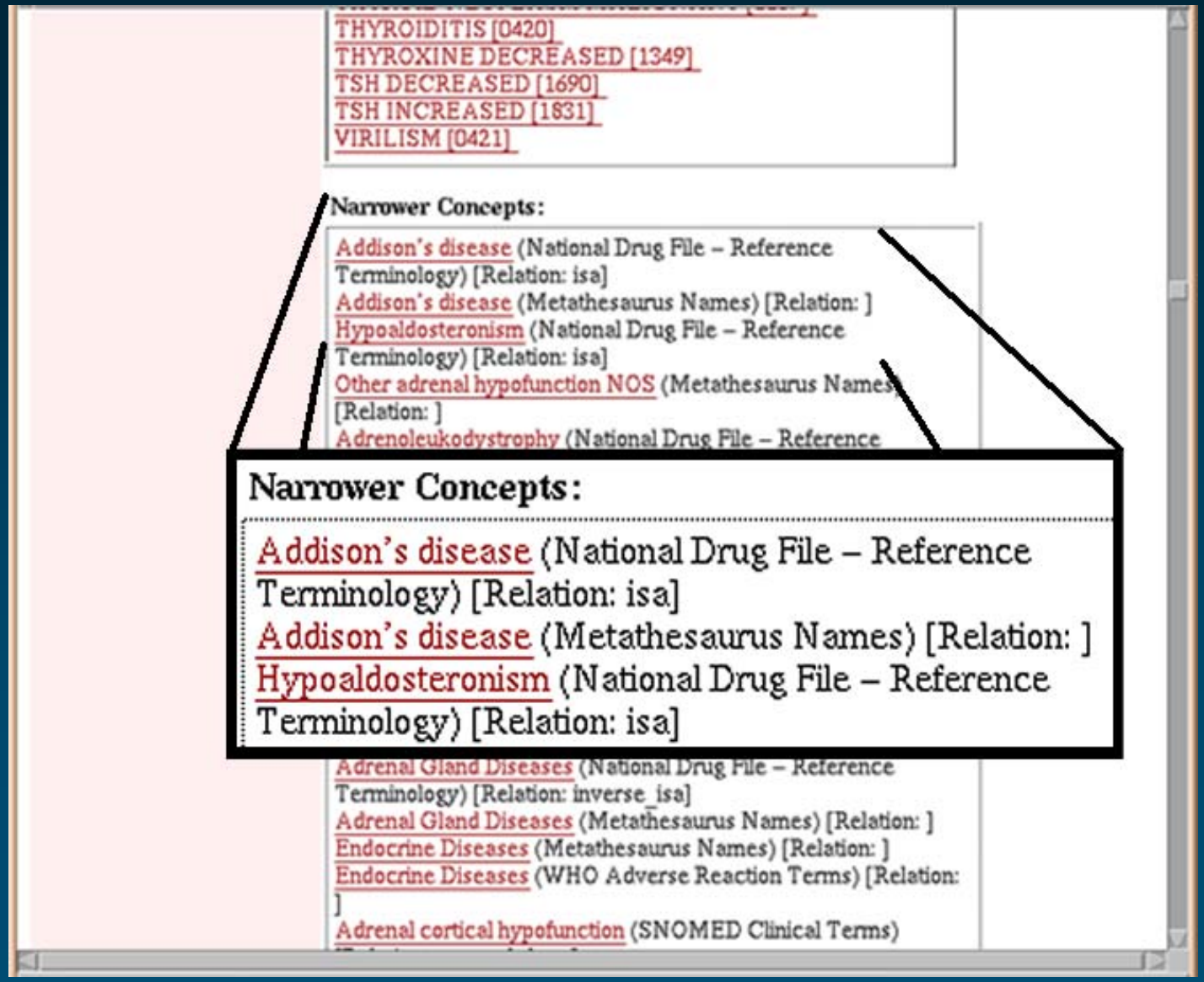
Concept Report Display All (continued)

Hierarchies



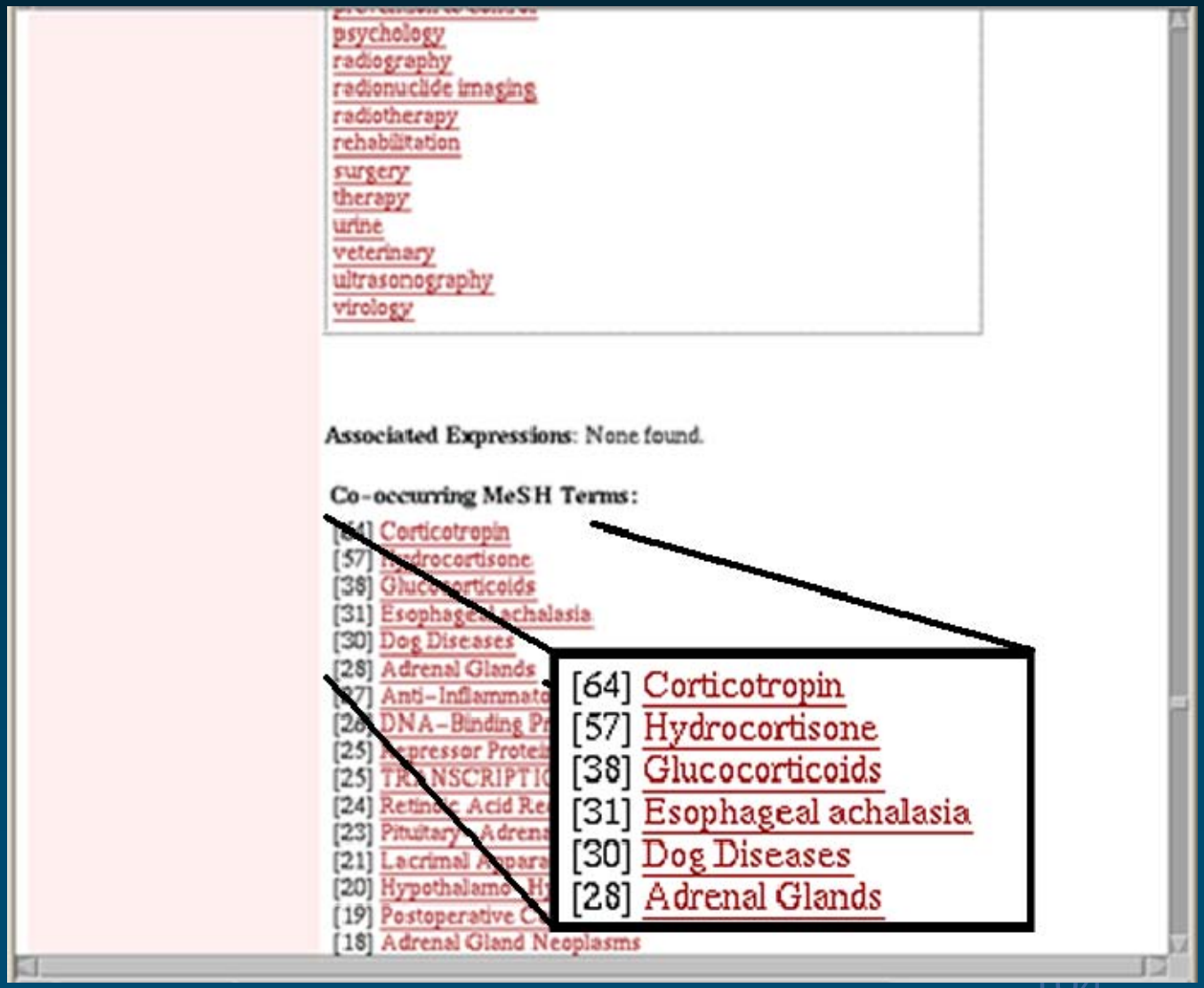
Concept Report Display All (continued)

Relations



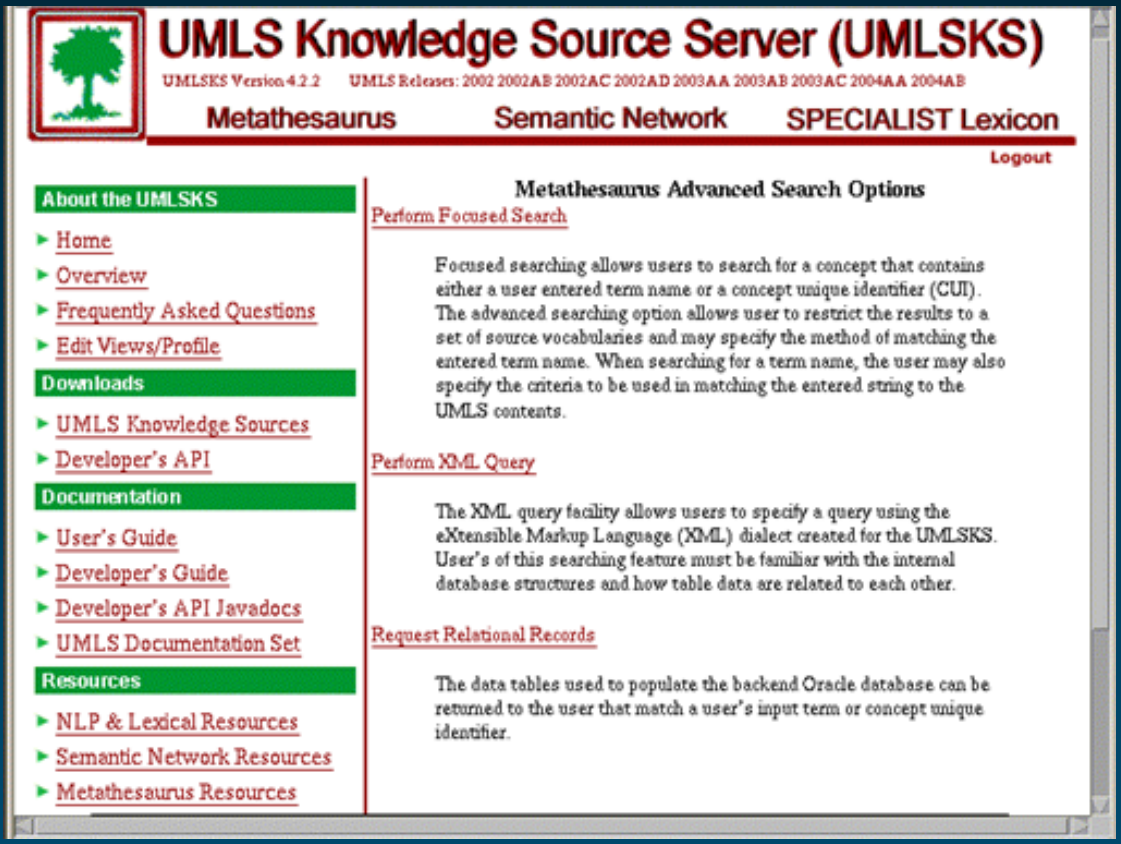
Concept Report Display All (continued)

Co-occurrence data



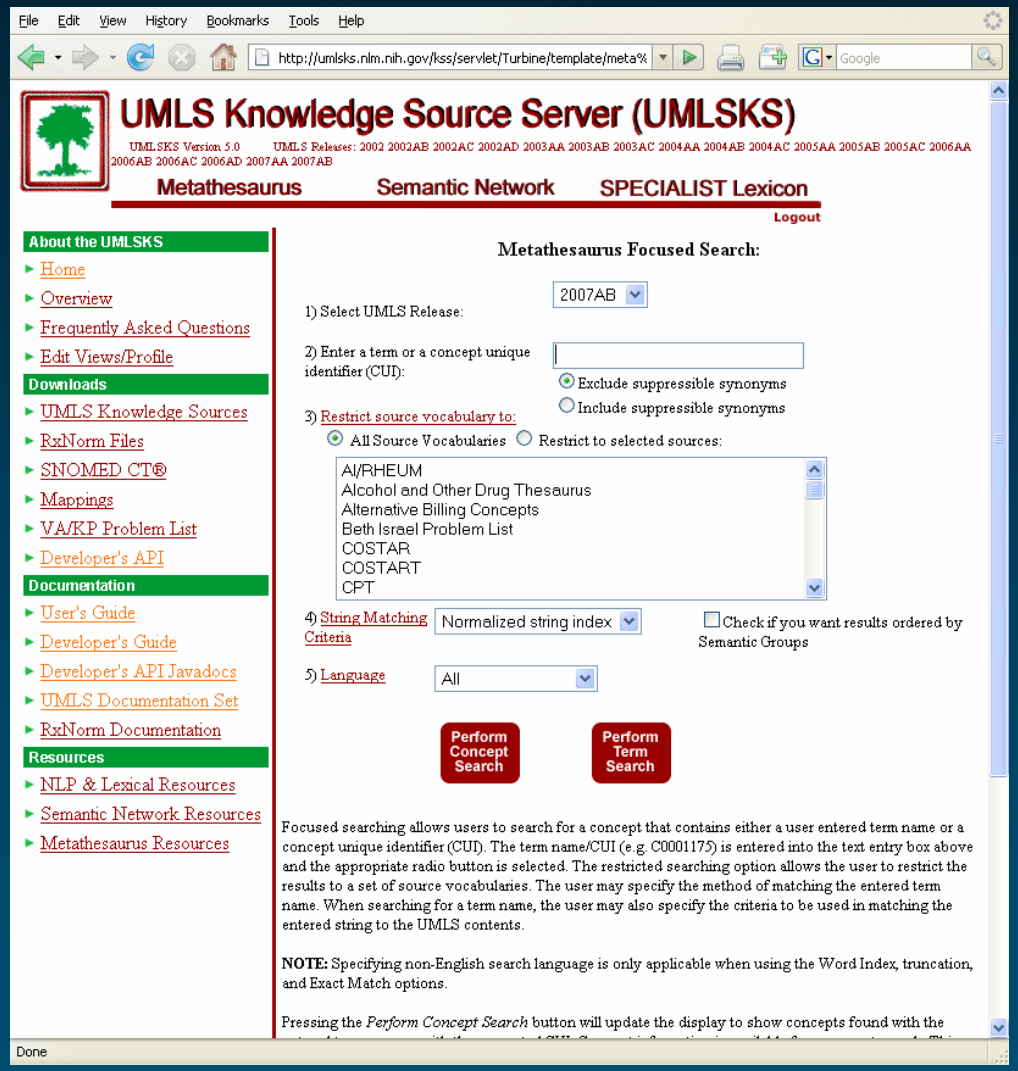
Metathesaurus Advanced Search Options

- Focused Search
- Raw Relational Records



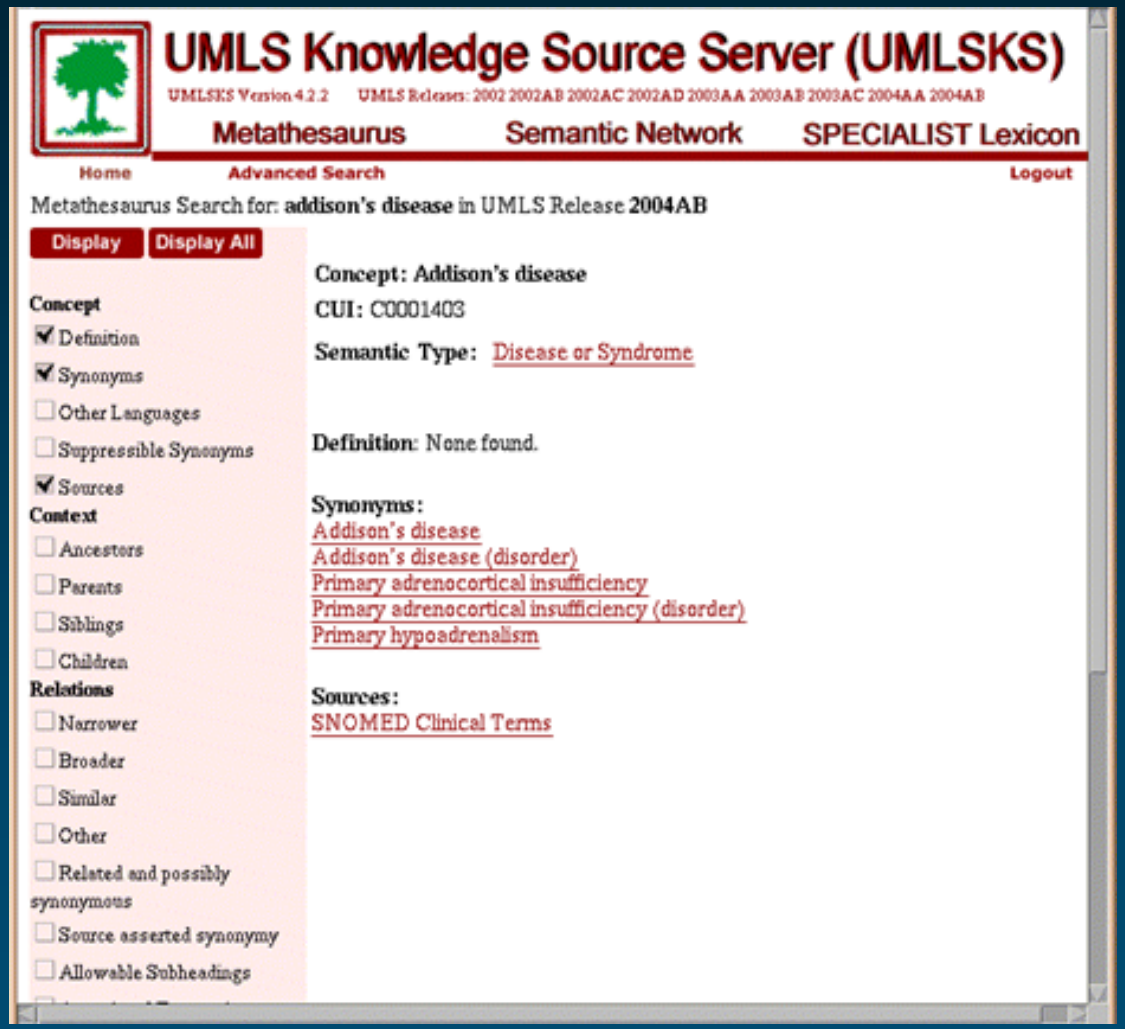
Metathesaurus Advanced Search Feature

Focused Search



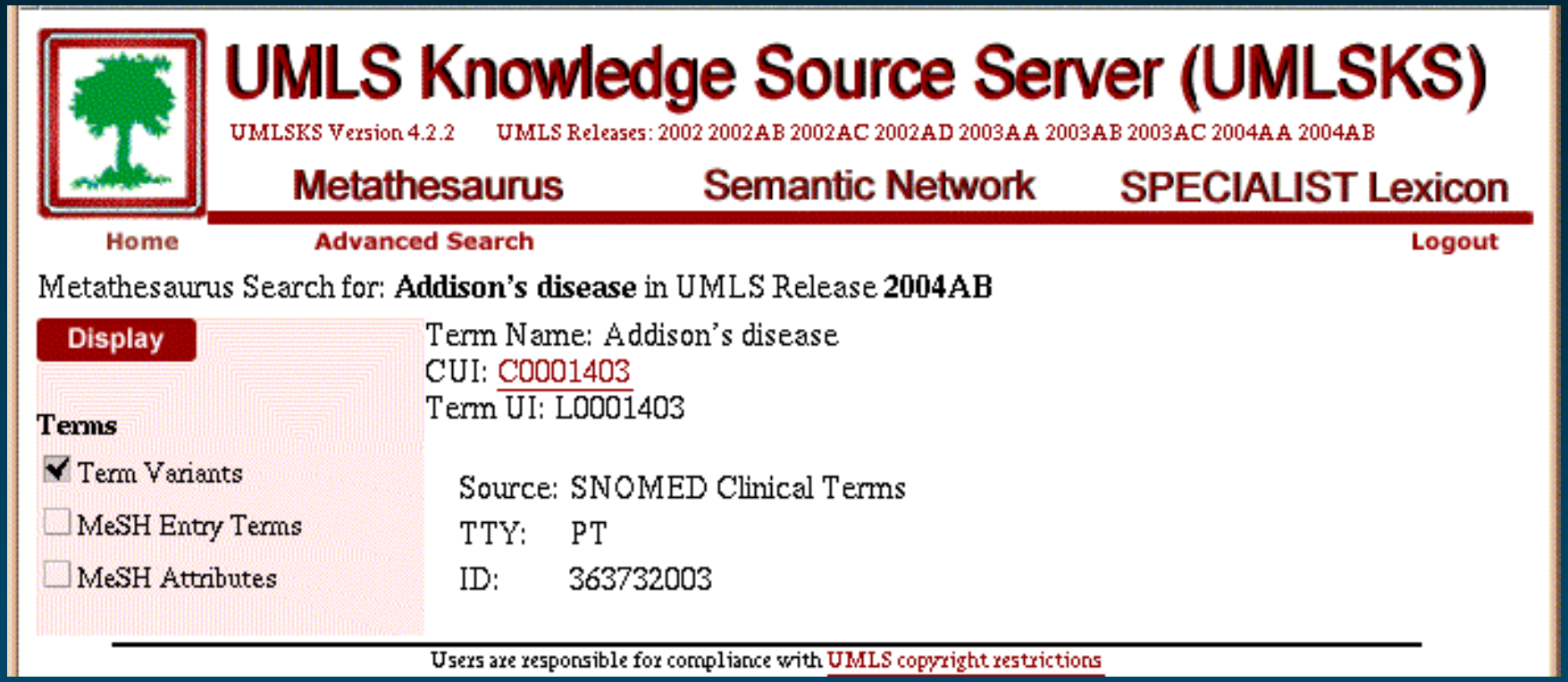
- UMLS Release
- Search Term
- Source Vocabularies
- String Criteria
  - Exact Match
  - Normalized string & word
  - Word
  - Truncation (left/right)
  - Approximate Match
- Language

Restricted Source Concept Report
Addison's Disease



- UMLS Release:
  2004AB
- Search Term:
  addison's disease
- Source Vocabulary:
  SNOMED CT
- String Criteria:
  Normalized string
- Language:
  English

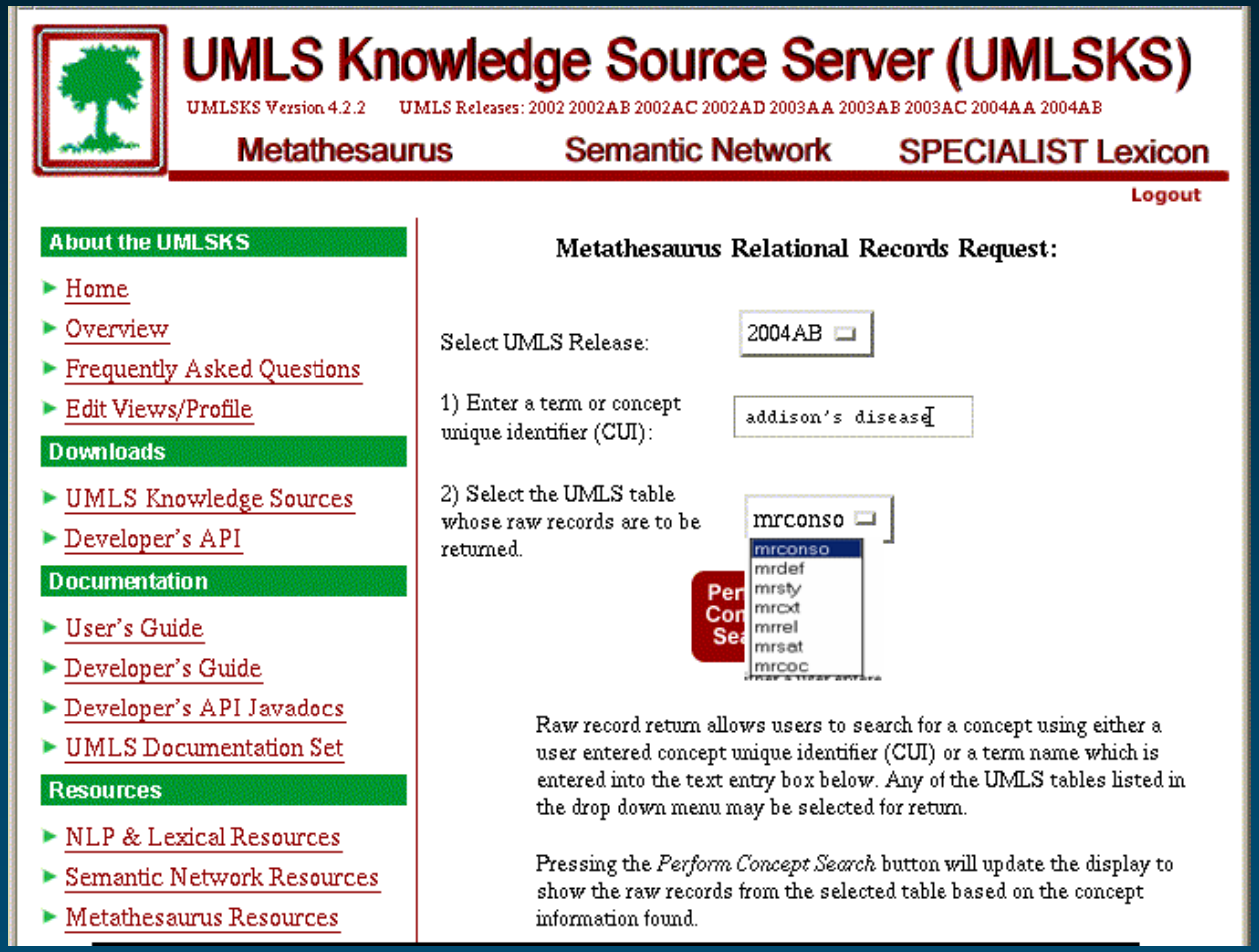
Addison's disease in SNOMED CT
Preferred Term and Code



- TTY: Term Type
- ID: Source Code Descriptor

Metathesaurus Advanced Search Feature
Relational Record Request

- UMLS Release
- Search Term
- UMLS Relational Table

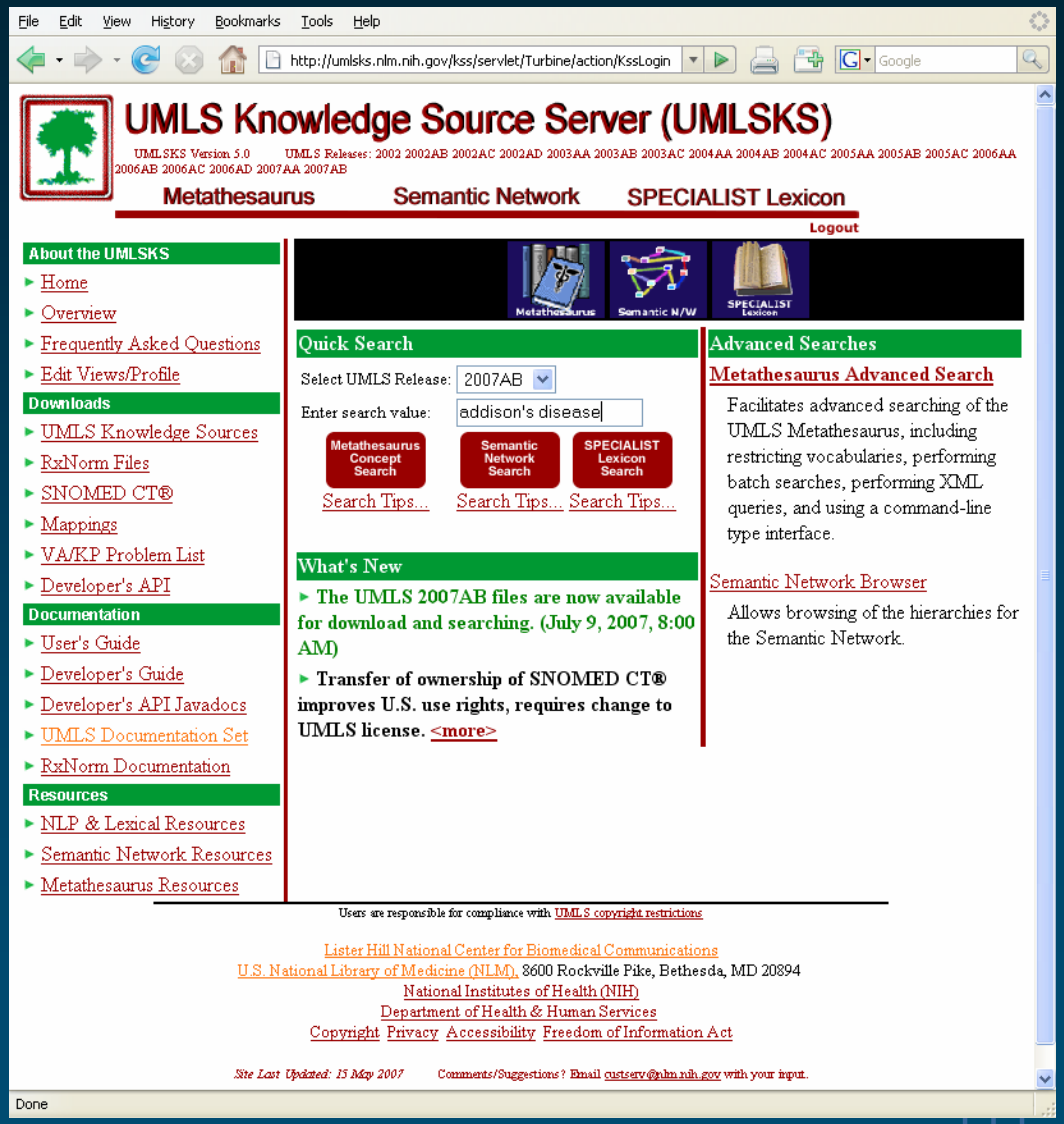


Relational Records MRCONSO.RRF

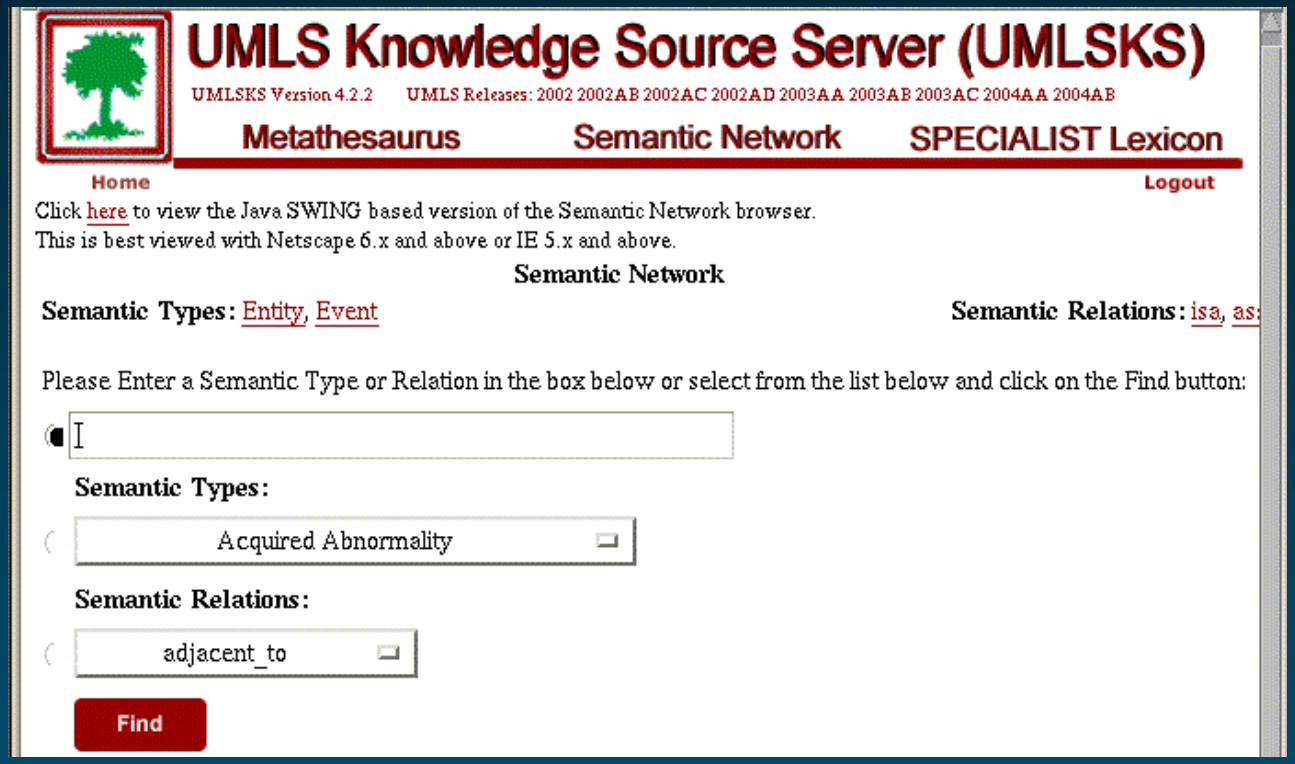
```
CUI|LAT|TS|LUI|STT|SUI|ISPREF|AUI|SAUI|SCUI|SDUI|SAB|TTY|CODE|STR|SUPPRESS|CVF|
C0001403|CZE|P|L3180742|PF|S3708232|Y|A3910108|||D000224|MSHCZE|MH|D000224|ADDISONOVA NEMOC|
C0001403|DUT|P|L2048638|PF|S2386860|N|A6566810|||10001130|MDRDUT|LT|10001130|Addison, ziekte
C0001403|DUT|P|L2048638|PF|S2386860|Y|A3931189|||D000224|MSHDUT|MH|D000224|Addison, ziekte v
C0001403|DUT|S|L2048637|PF|S2386859|Y|A3931188|||D000224|MSHDUT|SY|D000224|Addison, syndroom
C0001403|DUT|S|L2528364|PF|S2985131|Y|A3966882|||D000224|MSHDUT|SY|D000224|Ziekte van Addisc
C0001403|DUT|S|L3205108|PF|S3732602|N|A5146733||E27.1|||ICD10DUT|PT|E27.1|Primaire bijniersch
C0001403|DUT|S|L3205108|PF|S3732602|Y|A3970882|||D000224|MSHDUT|SY|D000224|Primaire bijniers
C0001403|DUT|S|L4999233|PF|S5686738|Y|A6627443|||10036696|MDRDUT|LT|10036696|primair hypoadr
C0001403|DUT|S|L4999270|PF|S5686775|N|A6627493|||10052381|MDRDUT|LT|10052381|primaire bijnie
C0001403|DUT|S|L4999270|PF|S5686775|Y|A6627494|||10052381|MDRDUT|PT|10052381|primaire bijnie
C0001403|DUT|S|L5012413|PF|S5699917|Y|A6645695|||10013096|MDRDUT|LT|10013096|ziekte van Addi
C0001403|ENG|P|L0001403|PF|S0354372|N|A0388276||||AOD|DE|0000006012|Addison's disease|0||
C0001403|ENG|P|L0001403|PF|S0354372|N|A0388277|||0060-3321|CSP|PT|0060-3321|Addison's diseas
C0001403|ENG|P|L0001403|PF|S0354372|N|A0388279||||LCH|PT|U000061|Addison's disease|0||
C0001403|ENG|P|L0001403|PF|S0354372|N|A0388280|||10001390|MDR|LT|10001130|Addison's disease|
C0001403|ENG|P|L0001403|PF|S0354372|N|A0388281||||RCD|PT|C1541|Addison's disease|3||
```

Semantic Network Searching

- Select Tab along top
- Quick search
- Advanced Search on right-hand side



Semantic Network Search



- Enter search string

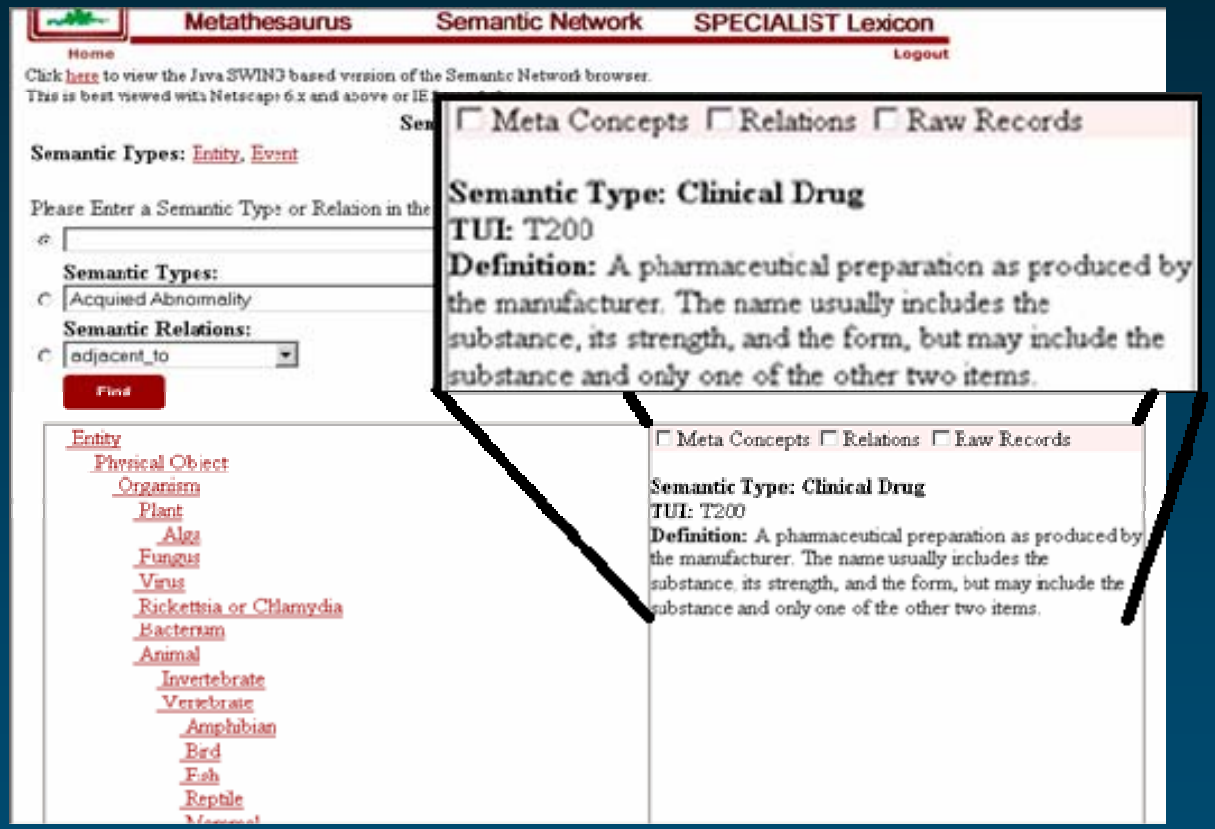
  -or-

- Select semantic type

  -or-

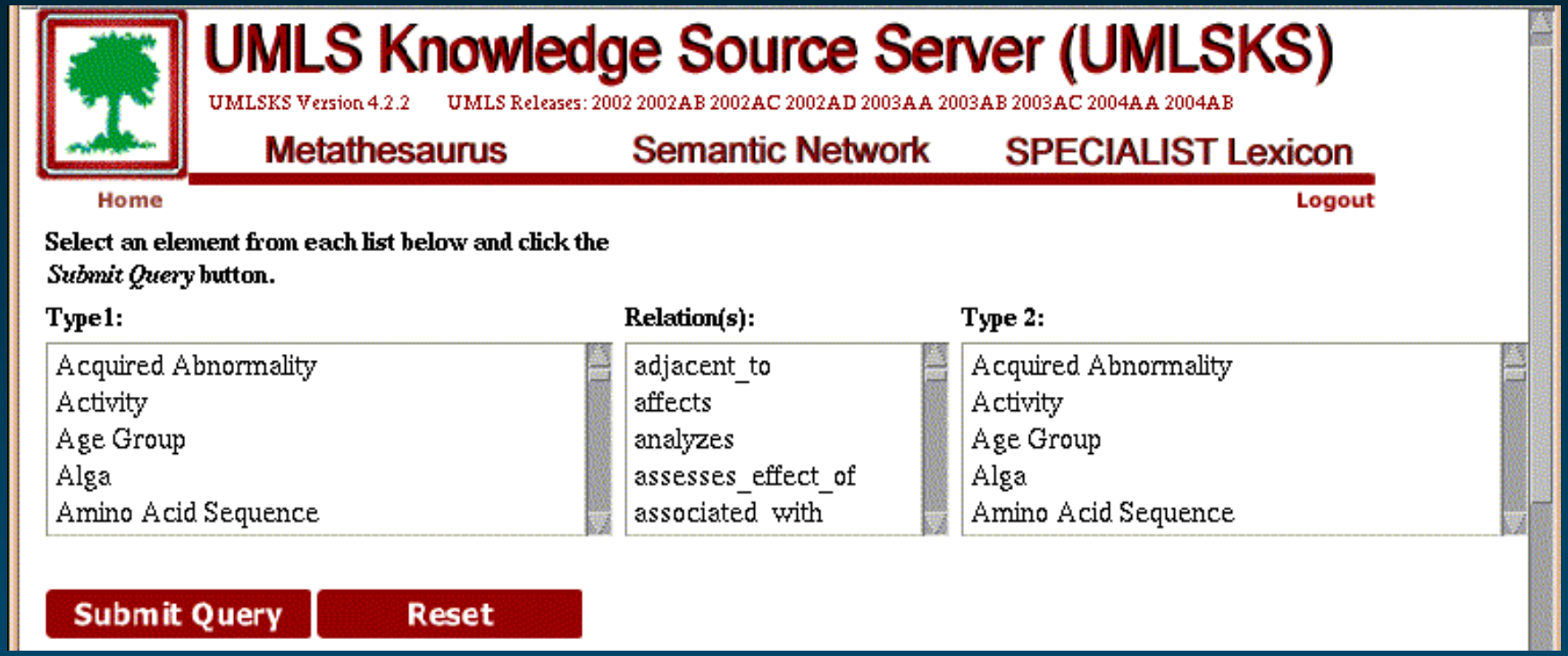
- Select semantic relation

Semantic Type Clinical Drug



- Browse ST hierarchy
- View Concepts with ST
- View Relations valid for the ST
- View Raw Relational Records

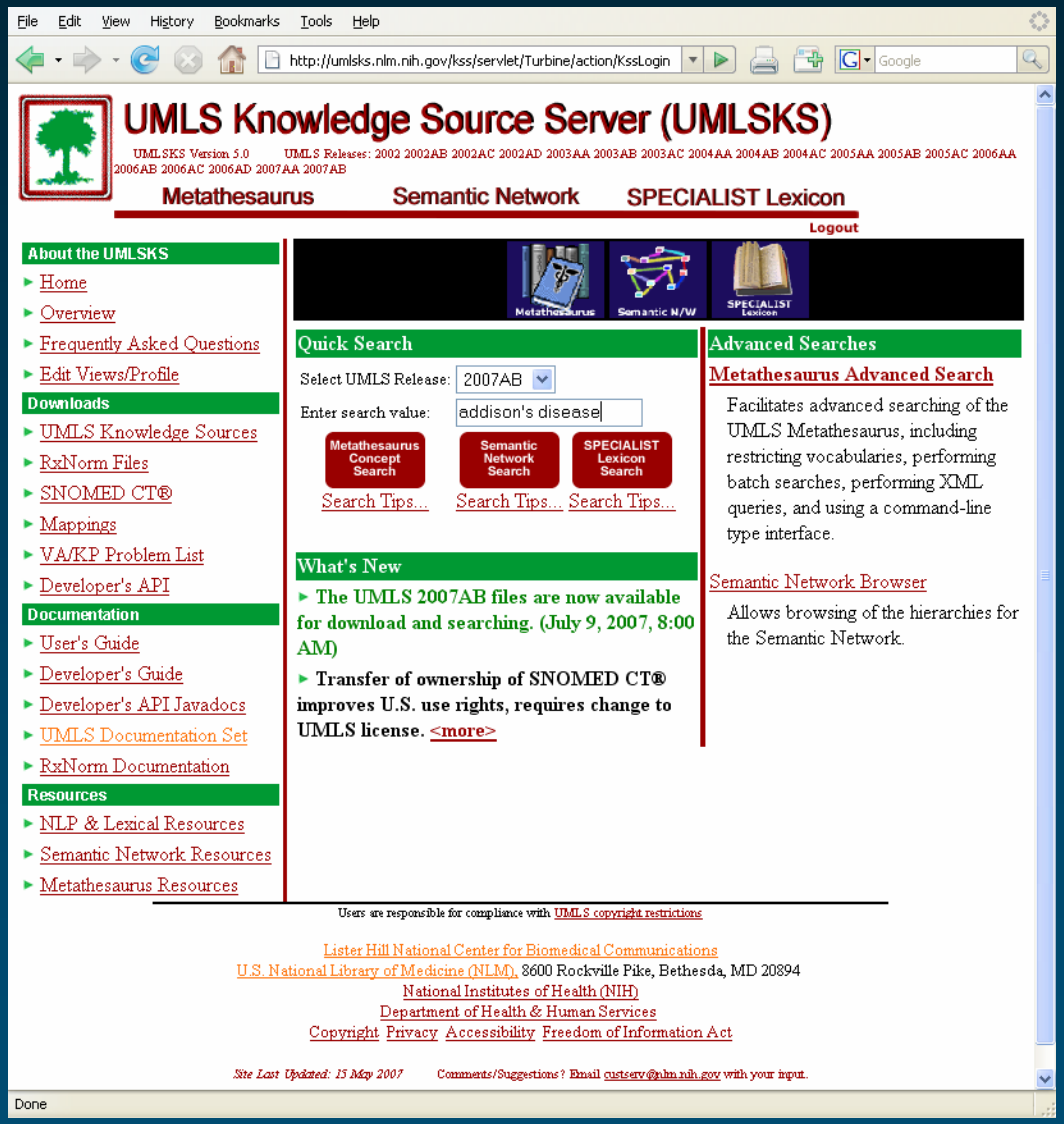
Show Relations Between Types



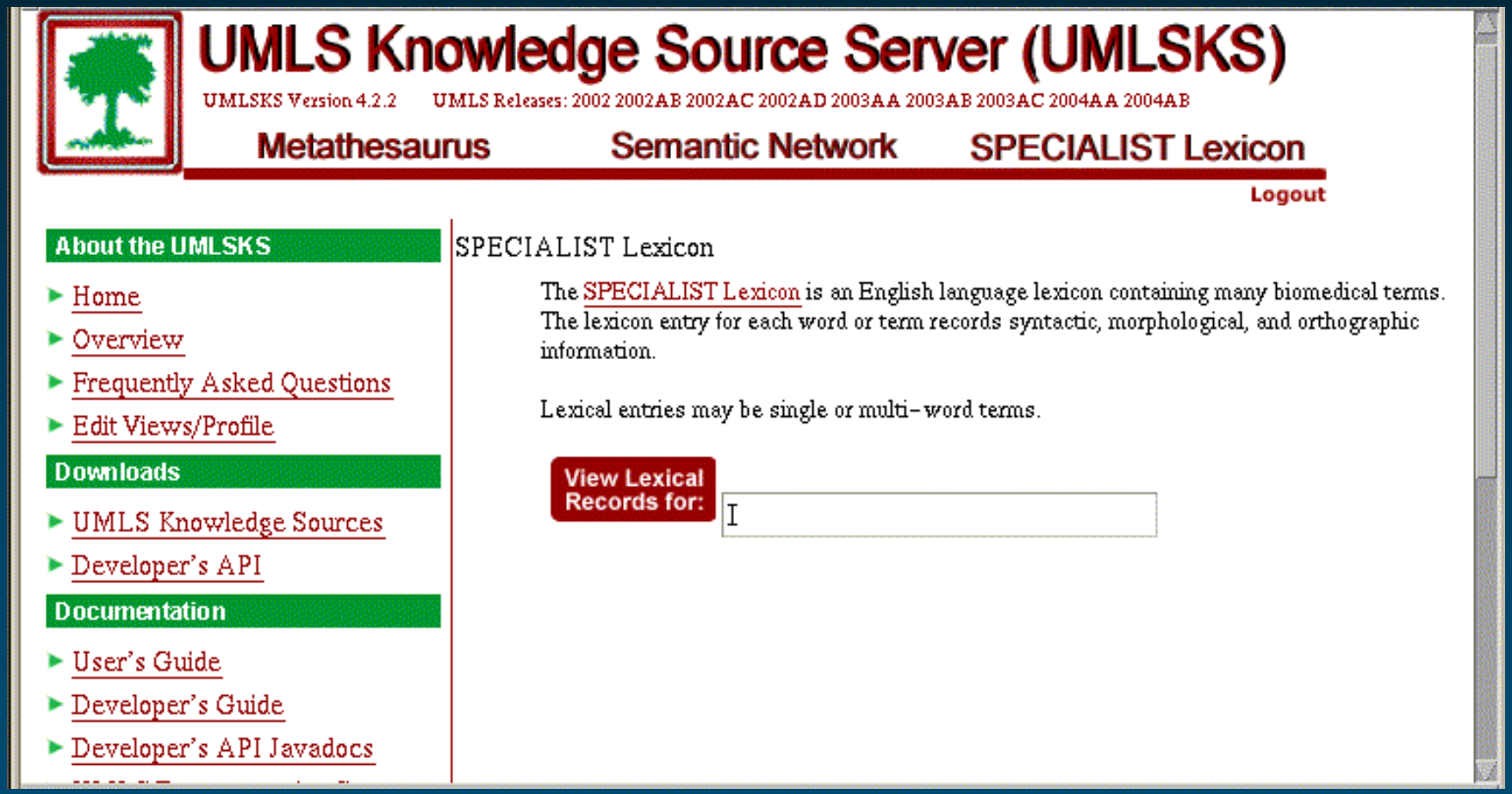
- Validates whether a selected Semantic Relationship (SR) holds between two selected Semantic Types (ST)

SPECIALIST Lexicon Searching

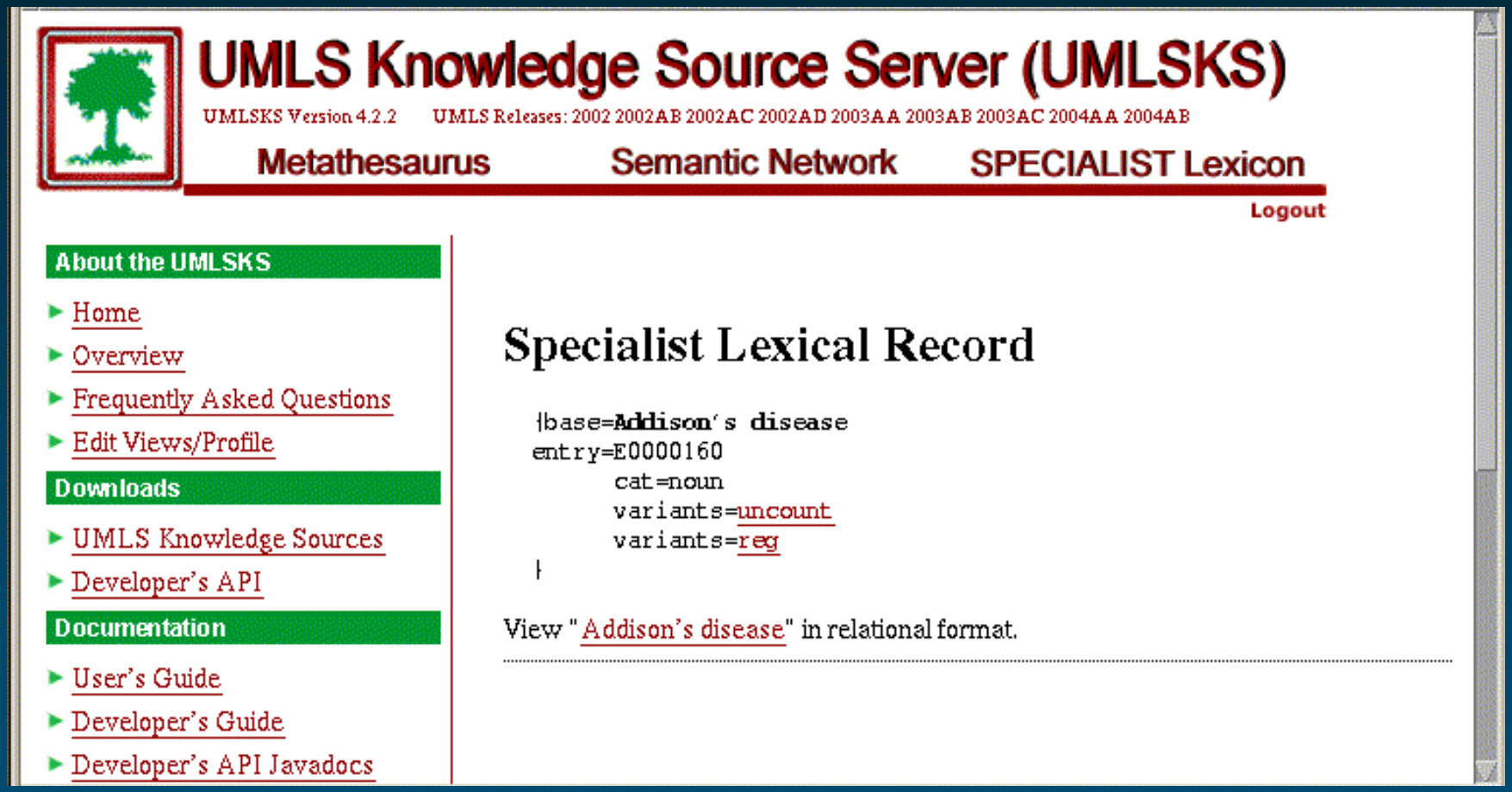
- Select Tab along top
- Quick search



SPECIALIST Lexicon Search

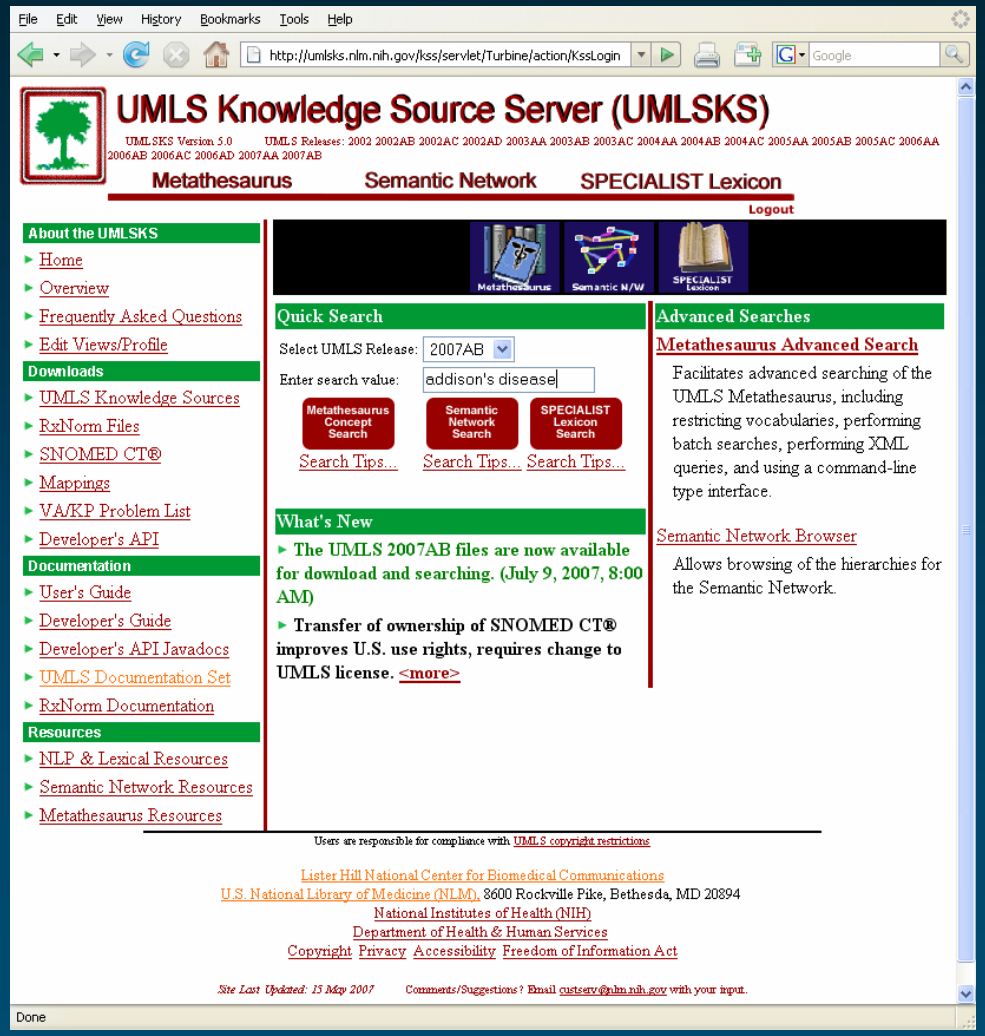


SPECIALIST Lexical Record



UMLS Resources

- NLP & Lexical Resources
  - MetaMap Transfer (MMTx)
  - Word Sense Disambiguation (WSD) Test Collection
- Semantic Network
  - Semantic Navigator
  - Semantic Groups
- Metathesaurus
  - String Properties



Resources

- NLP & Lexical Resources
- Semantic Network Resources
- Metathesaurus Resources

The “new” UMLSKS (coming soon)

Home | Help | Log Off

UMLS Knowledge Source Server (UMLSKS)

Home | Metathesaurus | SPECIALIST Lexicon | Semantic Network | My UMLSKS | Source View

Hello, Basic Umls-user

Downloads

- UMLS Knowledge Sources
- RxNorm Files
- Developer's API

Documentation

- Overview
- Frequently Asked Questions
- UMLSKS User's Guide
- Developer's API Javadocs
- UMLS Documentation
- RxNorm Documentation

Resources

- Lexical Tools
- MetaMap Transfer
- WSD Collection
- Semantic Navigator

Home | Change Content

Overview

Unified Medical Language System Knowledge Source Server (UMLSKS)
Fact Sheet Version 4.x

Background

The Unified Medical Language System® (UMLS) approach involves the development of a set of widely distributed Knowledge Sources (Metathesuarus, Semantic Network, and SPECIALIST Lexicon) that can be used by a variety of applications to compensate for differences in the way concepts are expressed in a variety of computerized biomedical sources.

The UMLS Knowledge Source Server (UMLSKS) is a computer application that provides Internet access to the Knowledge Sources and other related resources made available by developers using the UMLS. Its purpose is to make UMLS data more accessible to users, and in particular to system developers. The system architecture allows remote site users (individuals as well as computer programs) to send requests to a server at the National Library of Medicine (NLM). Access to the system is provided through the World Wide Web, an Extensible Markup Language (XML)-based socket programming interface, and through an Application Programmer Interface (API).

What's New

- **Web Services Implementation of the API.**
- **Web Interface based on Portal technology.**

Home | Help | Log Off

UMLS Knowledge Source Server (UMLSKS)

Home | Metathesaurus | SPECIALIST Lexicon | Semantic Network | My UMLSKS | Source View

Hello, Basic Umls-user

Choose a Section:

- Metathesaurus: Default View
- Metathesaurus: Lexical View
- Metathesaurus: Raw View

Downloads

- UMLS Knowledge Sources
- RxNorm Files
- Developer's API

Documentation

- Overview
- Frequently Asked Questions
- UMLSKS User's Guide
- Developer's API Javadocs
- UMLS Documentation
- RxNorm Documentation

Resources

- Lexical Tools
- MetaMap Transfer

Metathesaurus

Change

Metathesaurus Search

Enter term

Term: C0001403

OK

Release: 2007AA

Index: Exact Match

Sources: All sources

Basic Concept

Printer-Fri

Concept Information - UMLS Release 2007AA

[C0001403] Addison's disease

Definition:

CRISP Thesaurus/A0388277

MeSH/A6954527

An adrenal disease characterized by the progressive destruction of the ADRENAL CORTEX, resulting in insufficient production of ALDOSTERONE and HYDROCORTISONE. Clinical symptoms include ANOREXIA; NAUSEA; WEIGHT LOSS; MUSCLE WEAKNESS; and HYPERPIGMENTATION of the SKIN due to increase in circulating levels of ACTH precursor hormone which stimulates MELANOCYTES.

Semantic Types:

Disease or Syndrome

Atoms: (179)

Addison's disease[A0388276/Alcohol and Other Drug Thesaurus/DE/0000006012]

ADDISON DISEASE[A0385542/Clinical Problem Statements/PT/0022753]

ADRENAL INSUFFICIENCY (ADDISON'S DISEASE)[A0385630/COSTAR/PT/U000087]

Addison's disease[A0388277/CRISP Thesaurus/PT/0060-3321]

Home | Help | Log Off

UMLS Knowledge Source Server (UMLSKS)

Home | Metathesaurus | SPECIALIST Lexicon | Semantic Network | My UMLSKS | Source View

Hello, Basic Umls-user

Choose a Section:

- Metathesaurus: Default View
- Metathesaurus: Lexical View
- Metathesaurus: Raw View

Downloads

- UMLS Knowledge Sources
- RxNorm Files
- Developer's API

Documentation

- Overview
- Frequently Asked Questions
- UMLSKS User's Guide
- Developer's API Javadocs
- UMLS Documentation
- RxNorm Documentation

Resources

- Lexical Tools
- MetaMap Transfer
- WSD Collection

Metathesaurus

Change

Metathesaurus Search

Enter term

Term: C0001403

OK

Release: 2007AA

Index: Exact Match

Sources: MedDRA Japanese

Basic Concept

Printer-Friendly

Concept Information - UMLS Release 2007AA

[C0001403] Addison's disease

Semantic Types:

Atoms: (12)

- アジソン病[A11403670/MedDRA Japanese/LT/10001130]
- アジソン病[A11505319/MedDRA Japanese/OL/10013096]
- アジソン病[A11538038/MedDRA Japanese/PT/10001130]
- 原発性副腎機能低下症[A11549716/MedDRA Japanese/LT/10036696]
- 原発性副腎機能不全[A11513345/MedDRA Japanese/LT/10052381]
- 原発性副腎機能不全[A11520927/MedDRA Japanese/PT/10052381]
- ｹﾞﾝﾊﾟﾂｾｲﾌｸｼﾞﾝｷﾉｳﾌｾﾞﾝ[A11493725/MedDRA Japanese/LT/10052381]
- ｹﾞﾝﾊﾟﾂｾｲﾌｸｼﾞﾝｷﾉｳﾌｾﾞﾝ[A11500944/MedDRA Japanese/PT/10052381]
- ｱｼﾞｿﾝﾋﾞﾖｳ[A11443237/MedDRA Japanese/LT/10001130]
- ｱｼﾞｿﾝﾋﾞﾖｳ[A11525328/MedDRA Japanese/OL/10013096]
- ｱｼﾞｿﾝﾋﾞﾖｳ[A11538039/MedDRA Japanese/PT/10001130]
- ｹﾞﾝﾊﾟﾂｾｲﾌｸｼﾞﾝｷﾉｳﾃｲｶｼﾖｳ[A11529874/MedDRA Japanese/LT/10036696]

UMLS Semantic Navigator

- Web-based
  - http://mor.nlm.nih.gov/perl/semnav.pl
- Concept- and relation-centric
- Displays contexts graphically
- Displays all relations simultaneously
- Excludes hierarchical cycles in the UMLS graph
- Search
  - By CUI
  - By word

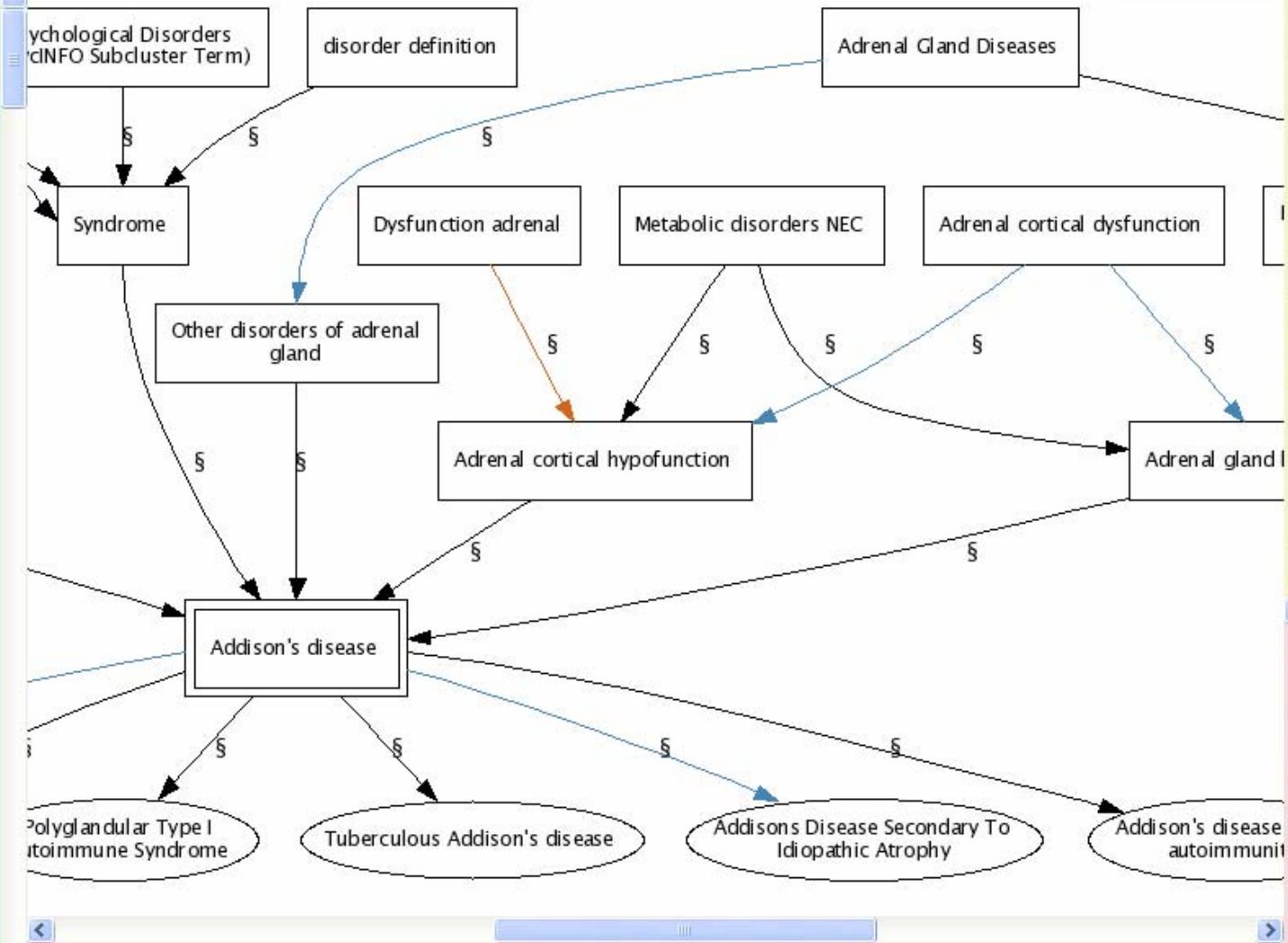
Siblings

Concepts & Ideas

- Clinical Syndromes

Disorders

- Aarskog syndrome
- Abstinence Syndrome
- Achard Syndrome
- Acroparesthesia Syndrome
- ACTH deficiency
- Acute coronary syndrome
- Addisonian crisis
- Adie Syndrome
- Adrenal cortex atrophy
- Adrenal insufficiency due to adrenal metastasis
- Adrenal insufficiency neonatal
- Adrenal Insufficiency Secondary To Hypopituitarism
- Adrenal medullary insufficiency



BCI

Addison's disease

LEGEND

Start again | Apply new parameters

Restrict to vocabulary: Show all

Highlight vocabulary: Nothing

UMLS data: UMLS_2007AA

Type of

Similar Concepts

- Adrenal cortical hypofunction

(1 concept)

Closest MeSH Terms

Main Headings

- Addison Disease

Other Related Concepts

Anatomy

- Abdominal Cavity
- Adrenal Cortex
- Adrenal Glands
- Endocrine system
- Retroperitoneal Space

Chemicals & Drugs

- Acetaminophen 25 MG/ML /

Co-occurring Concepts

Anatomy

- Adrenal Cortex [13]
- Adrenal Glands [16]
- Liver [3]
- Oral mucous membrane structure [2]

Chemicals & Drugs

- Alanine Transaminase [3]

Done

RRF Browser

- Distributed with the UMLS
  - Along with MetamorphoSys
- Standalone
- Can browse particular subsets of the Metathesaurus
- Search
  - By code
  - By CUI
  - By word

RRF Browser

Rich Release Format Browser 2007AB C0001403

File Edit Options Help

Monospaced 11 D:\2007AB\META

CUI Search | Code Search | Tree Browser | Word Search

Enter search terms: (English)

addison's disease

Search..

Select a concept. (3 results)

C0001403 Addison Disease
C0002892 Anemia, Pernicious
C0162309 Adrenoleukodystrophy

Raw View | Report View

Concept: [CUI C0001403] Addison Disease

Semantic Type

Disease or Syndrome

Definition

MSH/MH|An adrenal disease characterized by the progressive destruction of the ADRENAL CORTEX, resulting in insufficient production of ALDOSTERONE and HYDROCORTISONE. Clinical symptoms include ANOREXIA; NAUSEA; WEIGHT LOSS; MUSCLE WEAKNESS; and HYPERPIGMENTATION of the SKIN due to increase in circulating levels of ACTH precursor hormone which stimulates MELANOCYTES.

Atoms (17): (Sorted by Source, String)

ADDISON DIS [A12078968/MSH/DEV/D000224] Y
ADDISONS DIS [A12075312/MSH/DEV/D000224] Y
Addison Disease [A6954527/MSH/MH/D000224]
Addison's Disease [A6954528/MSH/EN/D000224]
Addisons Disease [A0019742/MSH/PM/D000224]
Adrenal Insufficiency, Primary [A6993206/MSH/PM/D000224]
Adrenocortical Insufficiencies, Primary [A6993209/MSH/PM/D000224]
Adrenocortical Insufficiency, Primary [A6993210/MSH/PM/D000224]
Disease, Addison [A0049628/MSH/PM/D000224]

Knowledge Source Server
Application Programming Interface

UMLSKS API basics

- Remote server at NLM
- Local application connected through

Java RMI

- Java-based applications
- Developer's Guide: Chapter 3
- Set of Java classes (part of the UMLSKS API download)
- Detailed *Javadoc* documentation online and with API download

TCP/IP socket

- XML-based queries
- Developer's Guide: Chapter 5
- XML schema
- Socket server
  - Host: umlsks.nlm.nih.gov
  - Port: 8042

Developer's Guide

About the UMLSKS
- Home
- Overview
- Frequently Asked Questions
- Edit Views/Profile

Downloads
- UMLS Knowledge Sources
- Developer's API

Documentation
- User's Guide
- Developer's Guide
  1. Introduction
  2. Installing the UMLSKS
  3. Building UMLSKS Software Applications
  4. Using the XML Query Facility
  5. Using the UMLSKS Socket Server
- UMLS Documentation Set

Resources
- NLP & Lexical Resources
- Semantic Network Resources
- Metathesaurus Resources

Developer's Guide

Print/Text Version | Table of Contents | About This Guide

About This Guide

This guide describes the installation and usage of the programmatic interface for the UMLS Knowledge Source Server (UMLSKS).

Audience

The audience for this guide is developers of UMLSKS applications using the UMLSKS API.

Release Notes

Please refer to the Release Bulletin for a detailed list of features, bug fixes, and known problems with this version of the UMLSKS.

How to Use This Guide

This manual contains the following chapters:

- Chapter 1 - Introduction describes the basic features and architecture of the UMLSKS.
- Chapter 2 - Installing the UMLSKS provides administrators instructions on installing and tailoring a UMLSKS installation.
- Chapter 3 - Building UMLSKS Software Applications describes the functions available to developers wanting to interface to the UMLSKS through another Java program.
- Chapter 4 - Using the XML Query Facility describes how to use the querying facility of the UMLSKS wherein users build XML queries to be executed.
- Chapter 5 - Using the UMLSKS Socket Server describes how to use the socket server to pass XML formatted commands or command-line type queries (e.g. ks -meta -c aids) that are to be executed by the server, with the results passed back to a client listening to the socket.

Documentation Java API

UMLS Knowledge Source Server (UMLSKS)

UMLSKS Version 5.0 UMLS Releases: 2002 2002AB 2002AC 2002AD 2003AA 2003AB 2003AC 2004AA 2004AB 2004AC 2005AA 2005AB 2005AC 2006AA 2006AB 2006AC 2006AD 2007AA 2007AB

U.S. National Library of Medicine
Lister Hill National Center for Biomedical Communications (LHNCBC)

About the UMLSKS
- Home
- Overview
- Frequently Asked Questions
- Edit Views/Profile

Downloads
- UMLS Knowledge Sources
- Developer's API

Documentation
- User's Guide
- Developer's Guide
  1. Introduction
  2. Installing the UMLSKS
  3. Building UMLSKS Software Applications
  4. Using the XML Query Facility
  5. Using the UMLSKS Socket Server
- UMLS Documentation Set

Developer's Guide

Print/Text Version | Table of Contents | About This Guide

UMLSKS API Download

The following instructions describe the procedures for downloading and installing the UMLSKS API. The sections include

- Downloading the UMLSKS API
- Building the Example .java Files
- Running the Client
- Running the ExpertClient
- Running the SocketClient
- Running the StandardQueryClient
- Available Documentation
- Sample Output and XML Query Examples

Downloading the UMLSKS API

The complete set of files for the UMLSKS API may be downloaded by clicking on the link below. Clicking the

NLM

Documentation Javadocs

All Classes

Packages

gov.nih.nlm.kss.api
gov.nih.nlm.kss.mode
gov.nih.nlm.kss.mode
gov.nih.nlm.kss.mode
gov.nih.nlm.kss.mode
gov.nih.nlm.kss.mode
gov.nih.nlm.kss.mode

All Classes

AdjEntry
AdvEntry
AssocExprsExecutor
AssociatedExp
AssociatedExpVector
AssociativeRelExister
AssociativeRelation
AssociativeRelationVe
Attr
AttrVector
AttributeContext
AttributeValue
AuxEntry
BasicConceptPropsEx
COContext
CUIExecutor
CatEntry
CompareUtil
Concept
ConceptAttribute
ConceptAttributeVect
ConceptAttrsExecutor
ConceptDelta
ConceptId
ConceptIdVector
ConceptNameExecutc
ConceptPropsExecutc
ConceptVector
ConstantSettings
Context

Overview Package Class Use **Tree** **Deprecated** **Index** **Help**

PREV NEXT FRAMES NO FRAMES

Packages

gov.nih.nlm.kss.api	
gov.nih.nlm.kss.models	
gov.nih.nlm.kss.models.lex	
gov.nih.nlm.kss.models.meta	
gov.nih.nlm.kss.models.meta.assocExp	
gov.nih.nlm.kss.models.meta.attribute	
gov.nih.nlm.kss.models.meta.concept	
gov.nih.nlm.kss.models.meta.context	
gov.nih.nlm.kss.models.meta.cooccurrence	
gov.nih.nlm.kss.models.meta.deltas	
gov.nih.nlm.kss.models.meta.locator	
gov.nih.nlm.kss.models.meta.meshentry	
gov.nih.nlm.kss.models.meta.relation	
gov.nih.nlm.kss.models.meta.source	
gov.nih.nlm.kss.models.sem	
gov.nih.nlm.kss.models.sem.rels	
gov.nih.nlm.kss.models.sem.units	
gov.nih.nlm.kss.query	
gov.nih.nlm.kss.query.lex	
gov.nih.nlm.kss.query.meta	
gov.nih.nlm.kss.query.sem	
gov.nih.nlm.kss.util	
gov.nih.nlm.nls.lexCheck.Lib	

Sample XML query (1) Current version

```
<?xml version="1.0"?>
<getCurrentUMLSVersion version="1.0"/>
```

```
<?xml version="1.0"?>
<CurrentUMLSYear version="1.0">
      2007AB
</CurrentUMLSYear>
```

Sample XML query (2) Concepts by string

```
<?xml version="1.0"?>
<findCUI version="1.0">
<conceptName>appendicectomy</conceptName>
<language>ENG</language>
<exact/>
<noSuppressibles/>
</findCUI>
```

```
<?xml version="1.0"?>
<ConceptIdCollection version="1.0">
  <release>2004AB</release>
   <conceptId>
    <cui>C0003611</cui>
    <cn>Appendectomy</cn>
   </conceptId>
</ConceptIdCollection>
```

Sample XML query (3) Concepts properties

```
<?xml version="1.0"?>
<getSemanticType version="1.0">
<cui>C0033572</cui>
</getSemanticType>
```

```
<?xml version="1.0"?>
<SemanticTypeCollection version="1.0">
<release>2004AB</release>
<cui>C0033572</cui>
<cn>Prostate</cn>
    <semanticType>
        <tui>T023</tui>
        <sty>Body Part, Organ,
            or Organ Component</sty>
    </semanticType>
</SemanticTypeCollection>
```

Sample XML query (4) Relationships

```
<?xml version="1.0"?>
<getRelations version="1.0">
<cui>C0033572</cui>
<rel>RO</rel>
</getRelations>
```

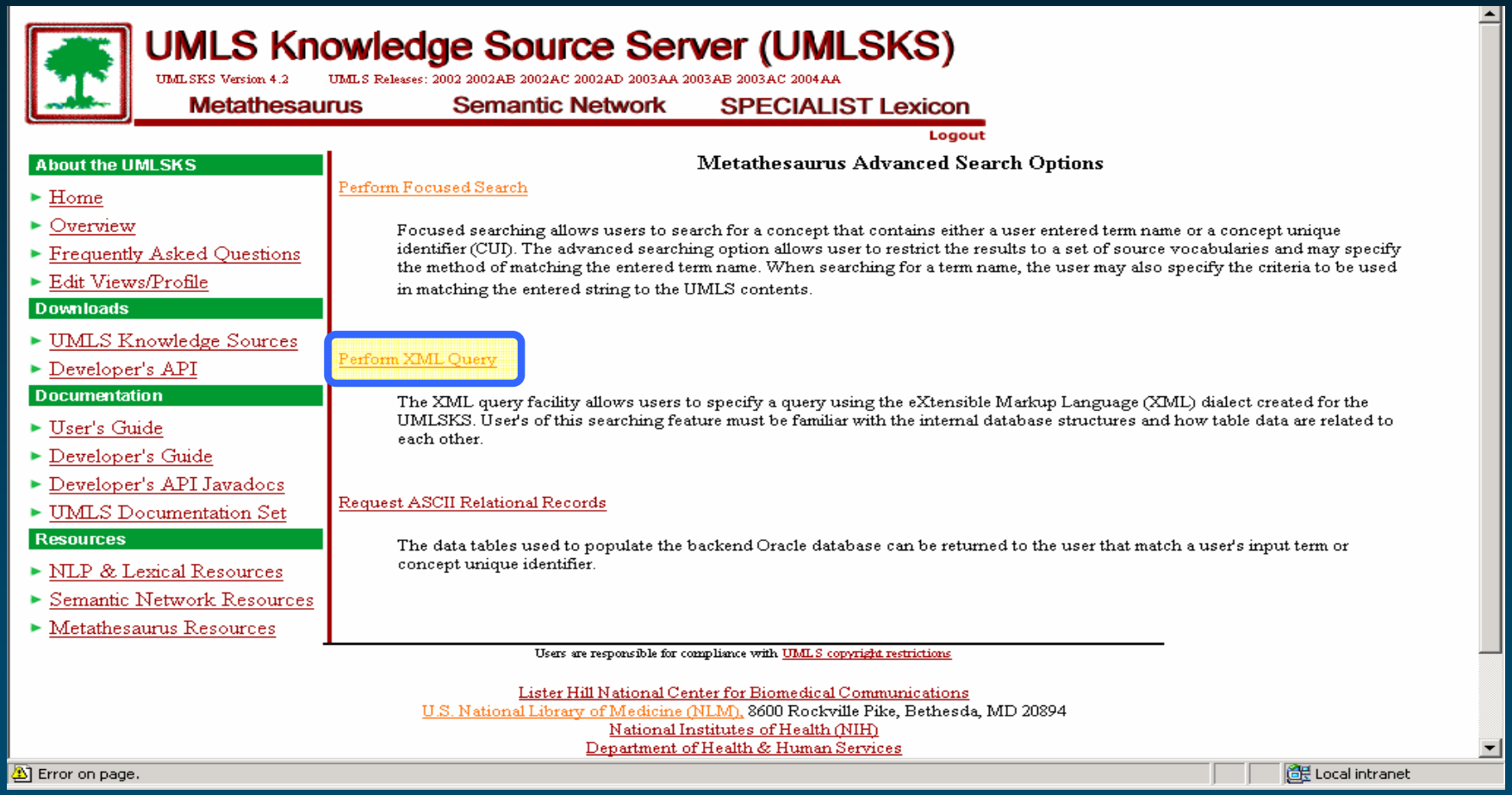
```
<?xml version="1.0"?>
<RelationCollection version="1.0">
[…]
   <relation>
     <aui>A3188910</aui>
     <sab>SNOMEDCT</sab>
     <relSource>
        <cui>C0007112</cui>
        <cn>Adenocarcinoma of prostate</cn>
        <aui>A3318222</aui>
        <rel>RO</rel>                    […]
        <rui>R54806623</rui>
        <rela>has_finding_site</rela>
      </relSource>
   </relation>
[…]
```

Sample XML query (5) All semantic type IDs

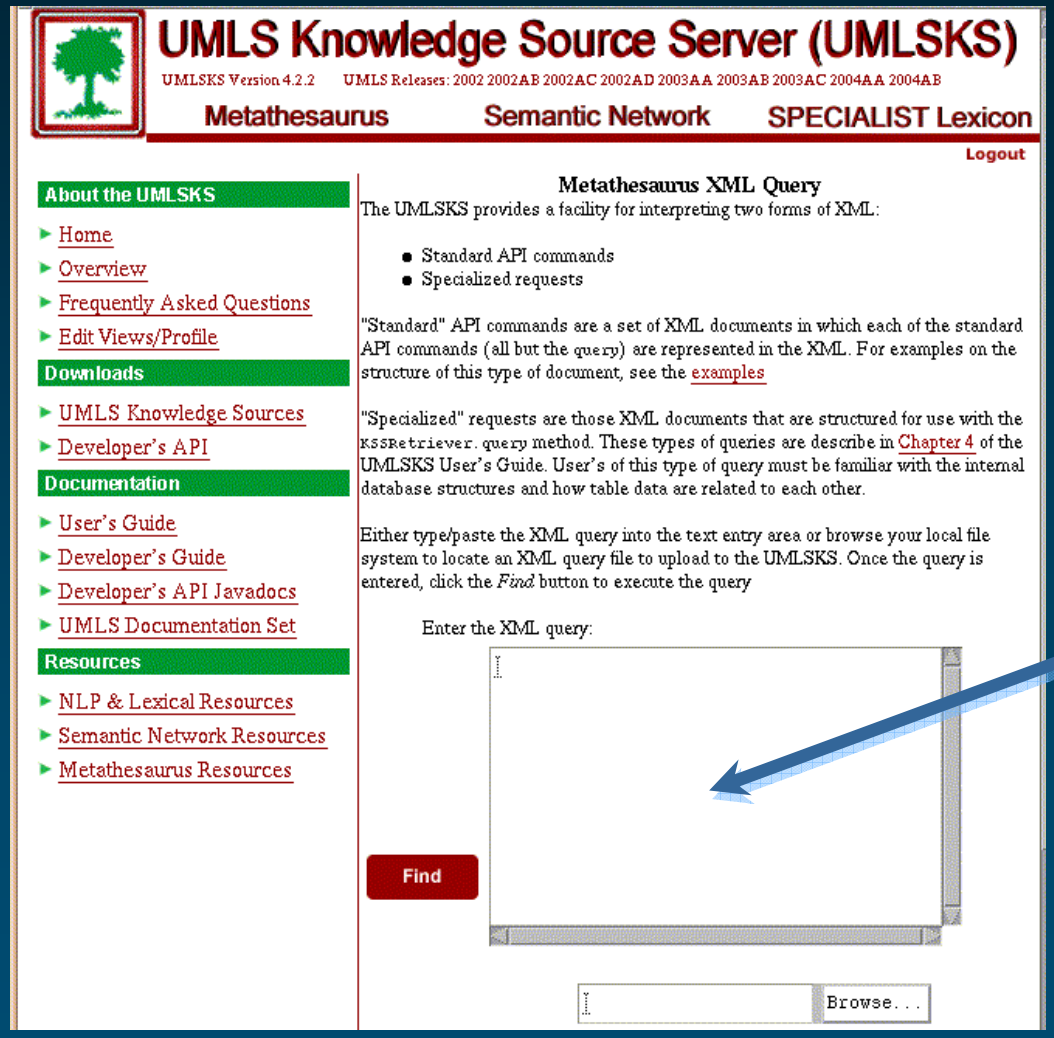
```
<?xml version="1.0"?>
<listSemTypeIds version="1.0">
</listSemTypeIds>
```

```
<?xml version="1.0"?>
<SemNetIdCollection version="1.0">
    <release>2004AB</release>
    <semnetId>
        <name>Acquired Abnormality</name>
        <ui>T020</ui>
        <semtype/>
    </semnetId>
    <semnetId>
        <name>Activity</name>
        <ui>T052</ui>
        <semtype/>
    </semnetId>
[…]
```

Performing XML queries from UMLSKS



Performing XML queries from UMLSKS




```
<?xml version="1.0"?>
<getRelations>
<cui>C0033572</cui>
<rel>RO</rel>
</getRelations>
```

Part II
How to use the UMLS?

(3) Installing the UMLS locally and Customizing the Metathesaurus using MetamorphoSys

What is MetamorphoSys?

- Tool distributed with the UMLS
- Multi-platform Java software
- The UMLS installation and customization wizard
  - Installs Knowledge Sources to local storage
  - Subsets and customizes a local Metathesaurus

Using MetamorphoSys

- Simple to use
- Screens and tabs lead you through process
- Installs NLM data format files to local storage

Why use MetamorphoSys?

Customize the Metathesaurus

- To remove terminology that is unhelpful, or even harmful, to your needs and purposes
- To comply with terms of license agreement

Why use MetamorphoSys?

Changing Default Settings

- To alter the preferred name
- To alter suppressibility of specific source term types

Customization is Critical

- Requires a clear understanding of:
  - Characteristics of source vocabularies
  - License arrangements
  - User's functional requirements
  - User's purpose and perspective
- Technical expertise

… and requires a multidisciplinary technical team

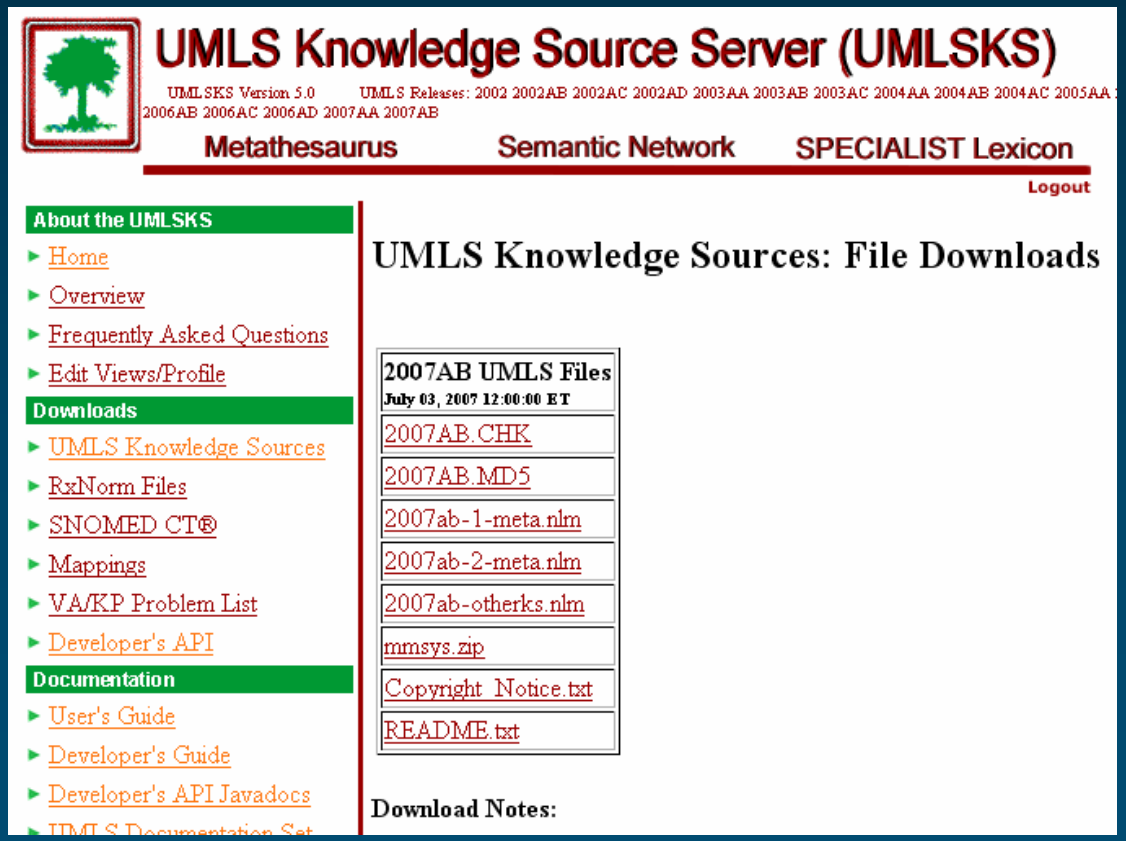
Machine Requirements

- A fast CPU – 1 GHz or higher
- 1 GB RAM recommended (512 MB min.)
- 6x (or better) DVD drive
- 22 GB minimum free disk space

- Runs on Sun Solaris 8 & 9, Windows XP, NT, and 2000, Linux, and Mac
- 1-10 hours run time on platforms tested

Download from UMLSKS …

- High speed Internet connection required
- Read the README file for the release

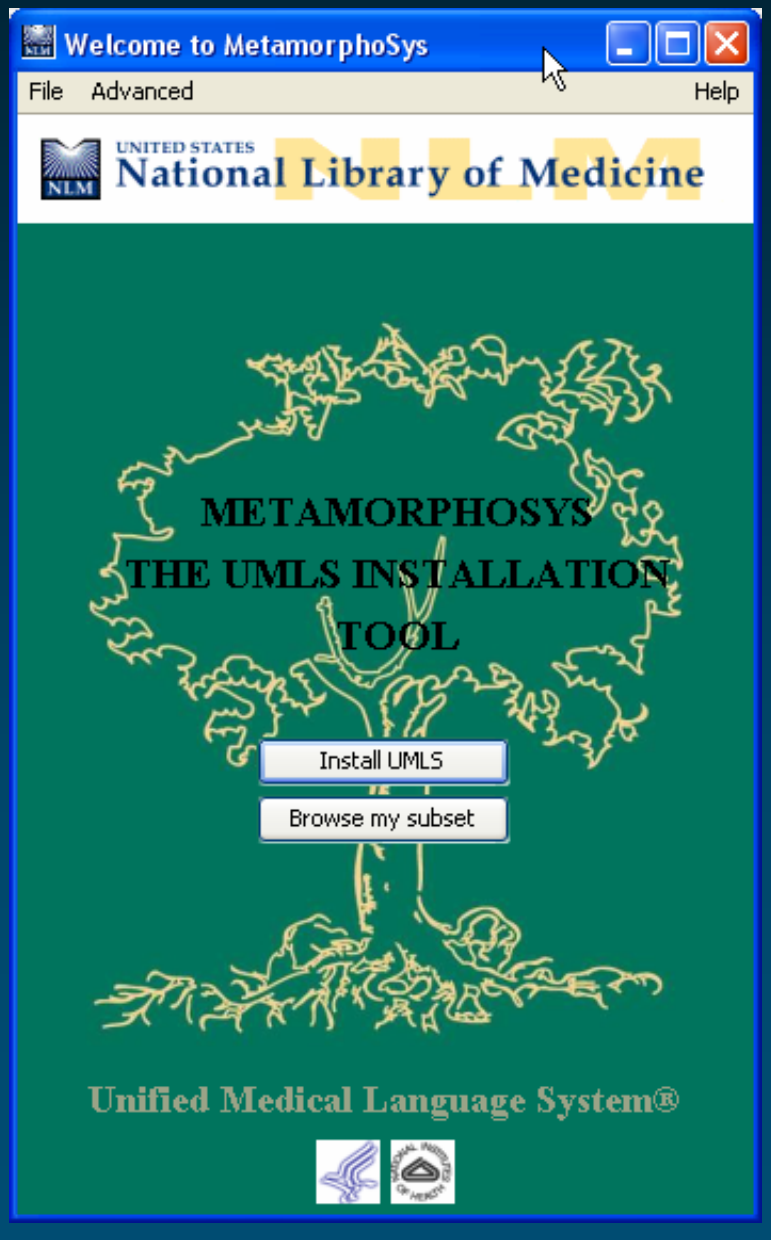


...or DVD?

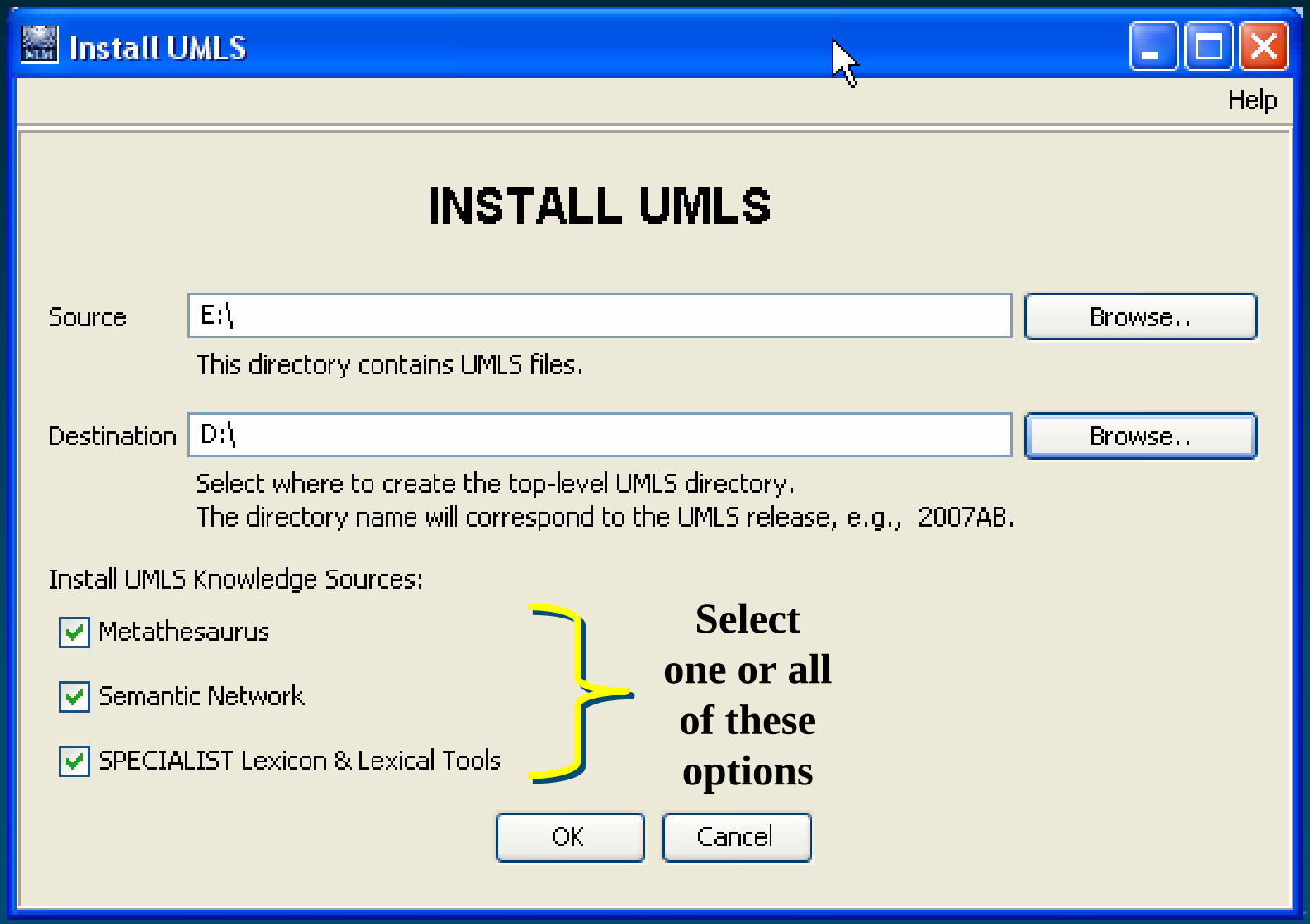
- Order at: umls_support@nlm.nih.gov
- **Include your license number**

- Run MetamorphoSys from DVD
  - Windows
    - Autorun; or go to root directory and click on “windows_mmsys.bat”
  - Linux, Solaris, Macintosh
    - open a terminal window, change to the root directory and type appropriate command: ./linux_mmsys.sh, ./solaris_mmsys.sh, ./macintosh_mmsys.sh

Welcome Screen



Install UMLS



UMLS License Notice

License Agreement Notice

Notice:

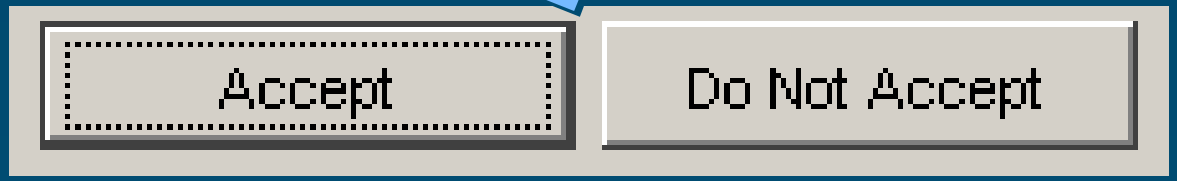
The Metathesaurus contains source vocabularies produced by many different copyright holders. The majority of the content of the Metathesaurus is available for use under the basic (and quite open) terms described in the Metathesaurus license http://www.nlm.nih.gov/research/umls/license.html

However, some vocabulary producers place ADDITIONAL RESTRICTIONS ON THE USE OF THEIR CONTENT AS DISTRIBUTED WITHIN THE METATHESAURUS.

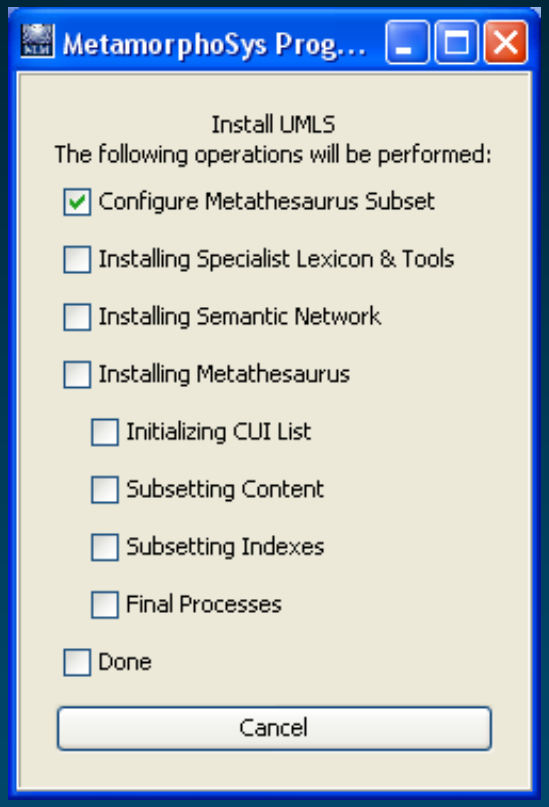
The various levels of additional restrictions are described in Section 12 of the license. The level that applies to individual vocabularies is recorded in the Appendix to the license. If a UMLS user already has a separate license for use of one of the source vocabularies, the user's existing license also applies to that source as distributed within the Metathesaurus. In some cases, UMLS users may have to request permission or negotiate a separate license with a vocabulary producer in order to use that vocabulary in a production system. There may be a charge associated with these separate permissions or license agreements.

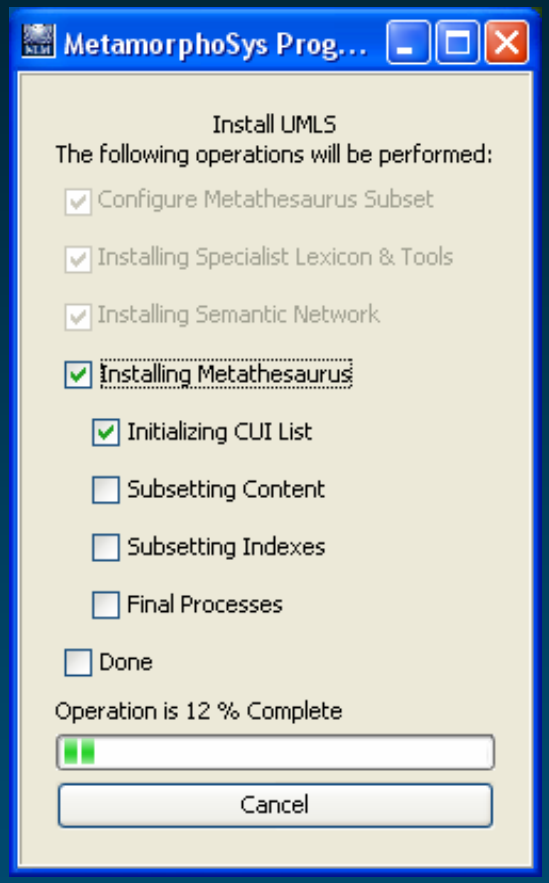
Please select "Accept" or "Do Not Accept" below after reviewing the license agreement at the URL above.

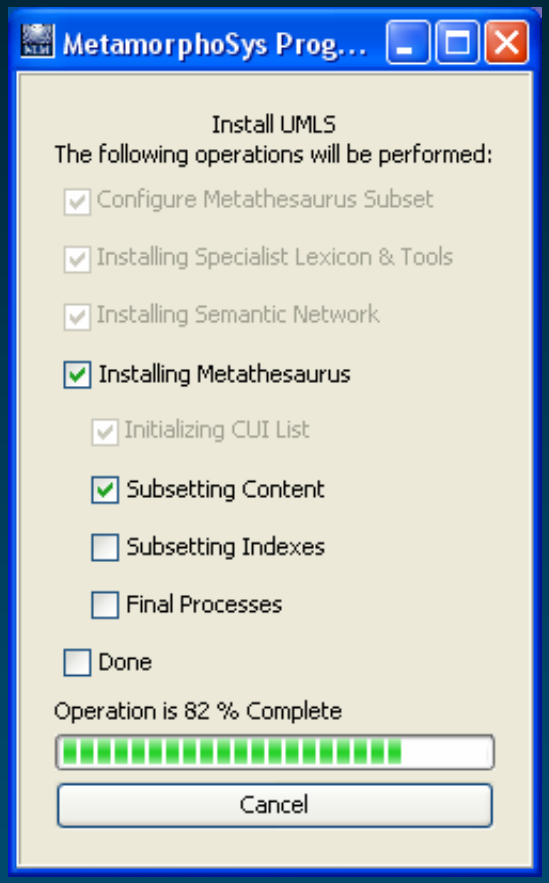
Accept | Do Not Accept



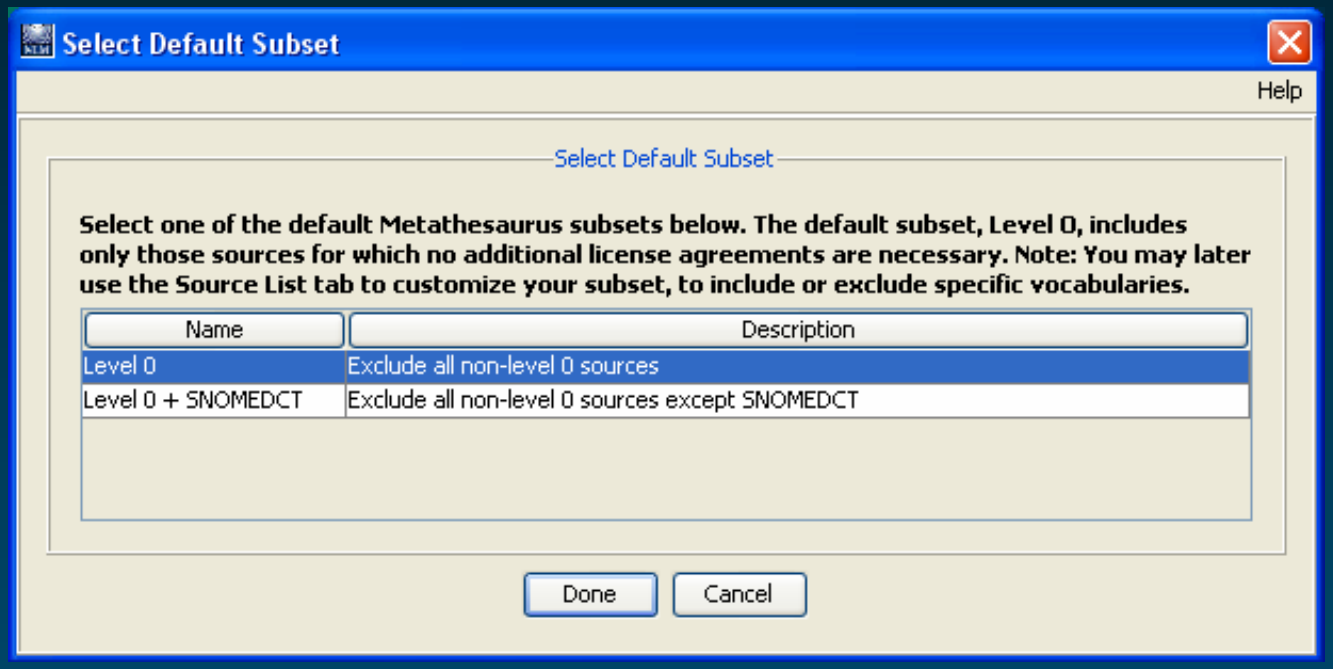
Installation progress monitor







Select a default subset



Level 0 → no separate additional license agreements

Level 0 + SNOMEDCT → Users from non-IHTSDO member countries must have separate license agreements

Input Options Tab

UMLS Metathesaurus Configuration 2007AB

File Edit Options Reset Done Help

Input Options | Output Options | Source List | Precedence | Suppressibility

Select the format in which subset data was saved and indicate the location of the source folder that contains the existing subset configuration. See Help for more information.

Users may proceed to the other Options tabs in any order by clicking on the tabs across the top of the screen or by clicking "Done" on the File bar and selecting "Begin Subset".

Input Format Options

Select Input Format

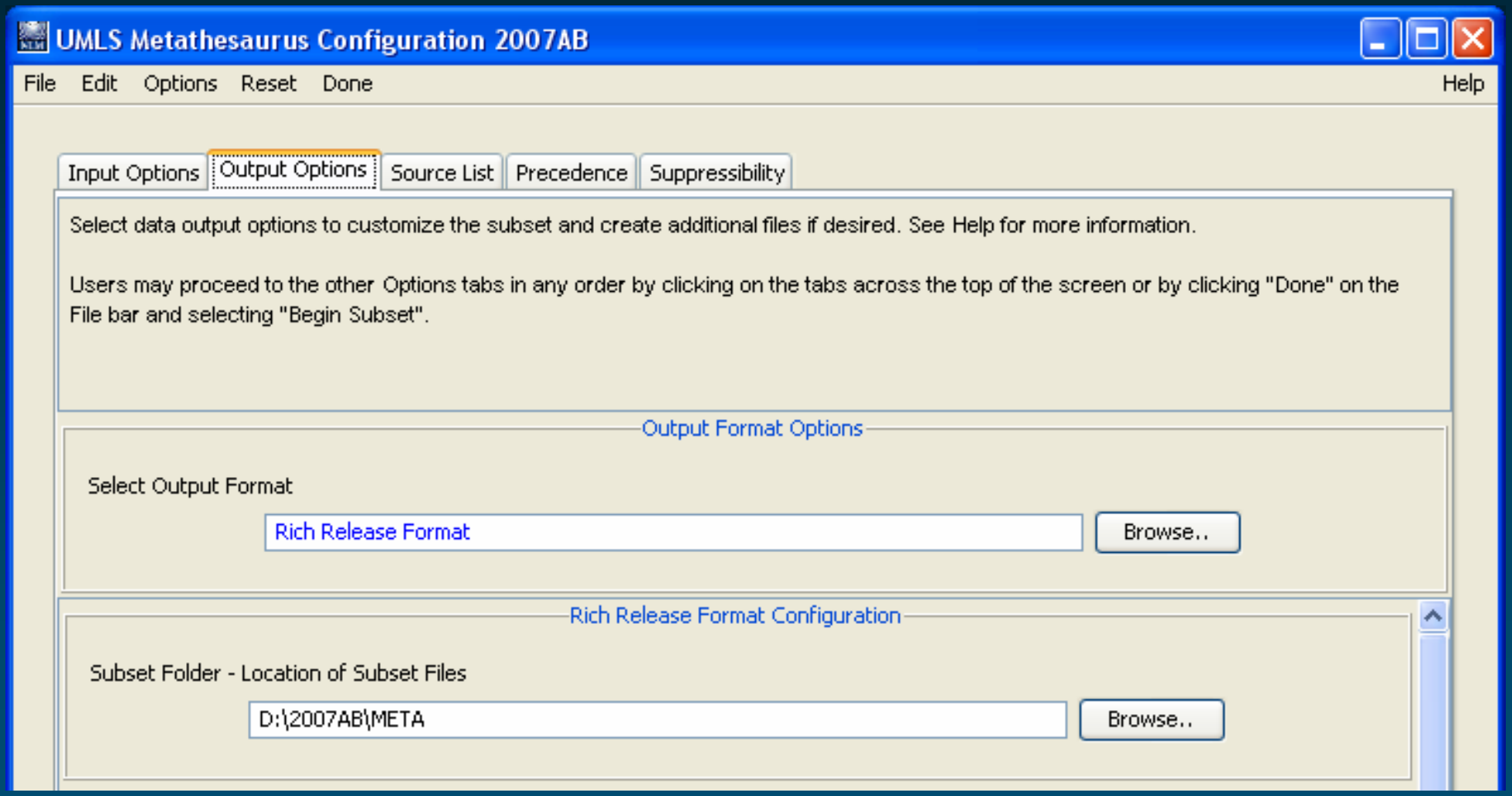
NLM Data File Format [Browse..]

NLM Data File Format Configuration

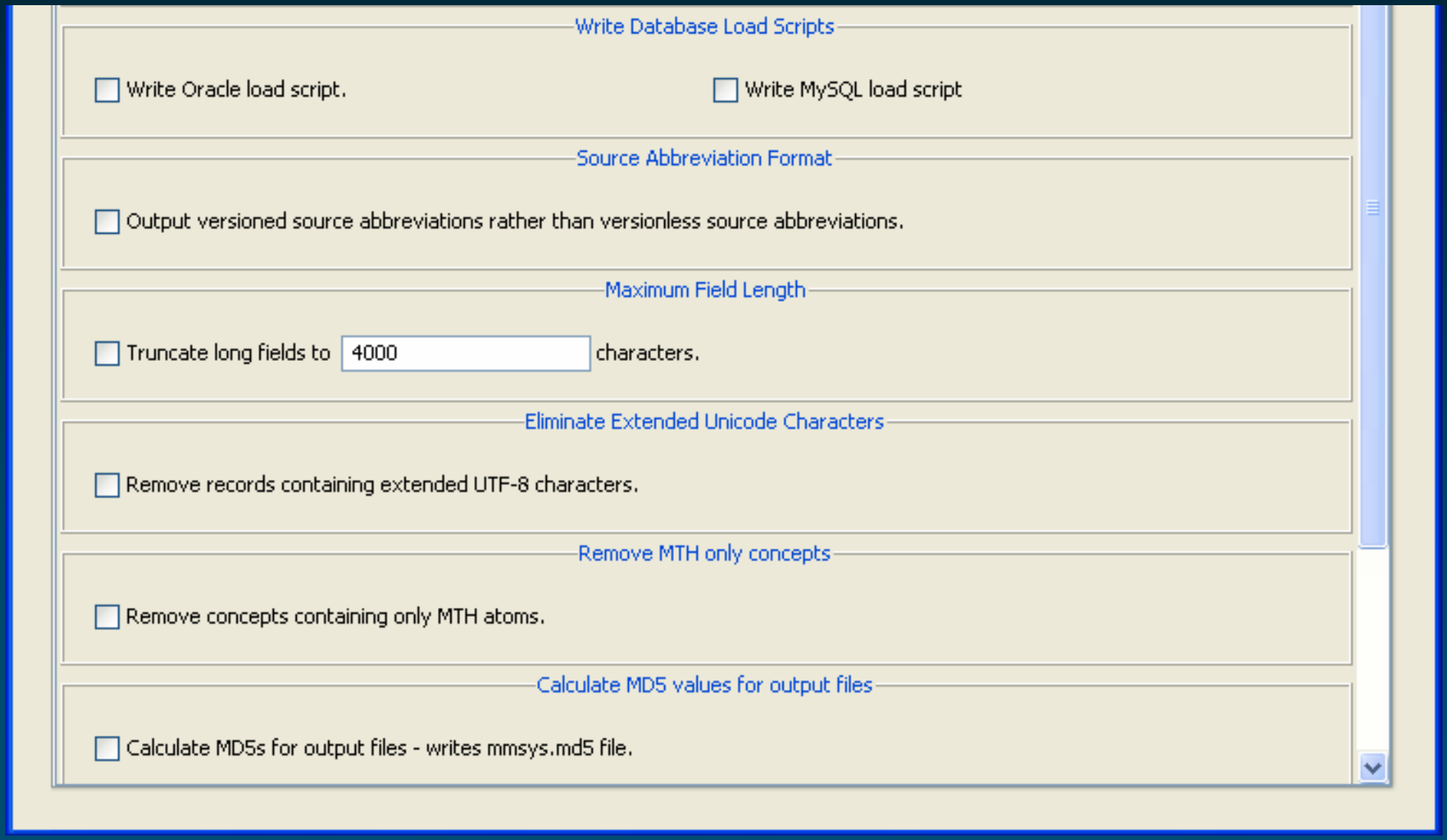
Source Folder - Location of Metathesaurus Files

E:\ [Browse..]

Output Options Tab



Output Options Tab



Source List Tab

UMLS Metathesaurus Configuration 2007AB

File Edit Options Reset Done Help

Input Options | Output Options | Source List | Precedence | Suppressibility

Indicate below to INCLUDE or EXCLUDE selected sources.
SELECTED sources appear on a dark background, e.g., AI/RHEUM, 1993.

To undo selections and return to default source list, select "Reset Source List" from Reset menu at top.

- (•) Select sources to EXCLUDE from subset
- ( ) Select sources to INCLUDE in subset

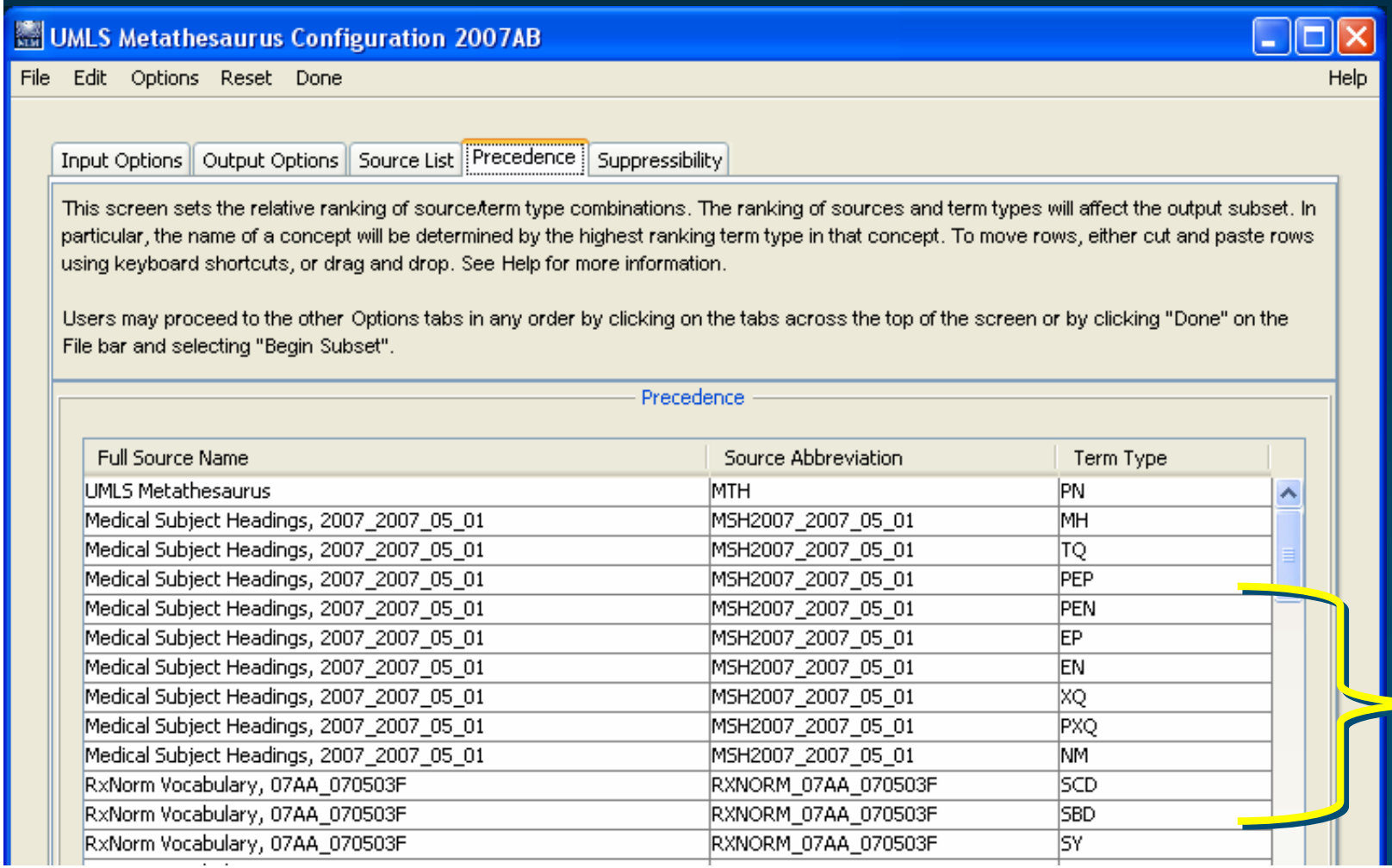
Sources to Exclude

*	Full Source Name	Source Abbreviation	Source Family	Lang...	Level	Concepts
	AI/RHEUM, 1993	AIR93	AIR	ENG	0	630
	Alternative Billing Concepts, 2006	ALT2006	ALT	ENG	3	4661
	Alcohol and Other Drug Thesaurus, 2000	AOD2000	AOD	ENG	0	15900
	Authorized Osteopathic Thesaurus, 2003	AOT2003	AOT	ENG	0	278
	Beth Israel Vocabulary, 1.0	BI98	BI	ENG	2	939
	Canonical Clinical Problem Statement System, 1999	CCPSS99	CCPSS	ENG	3	15266
	Clinical Classifications Software, 2005	CCS2005	CCS	ENG	0	1107
	Current Dental Terminology 2007-2008 (CDT-7/8), 2007-2008	CDT2007-2008	CDT	ENG	3	582
	COSTAR, 1989-1995	COSTAR_89-95	COSTAR	ENG	0	3073
	Medical Entities Dictionary, 2003	CPM2003	CPM	ENG	2	3078
	Physicians' Current Procedural Terminology, Spanish Transla...	CPT01SP	CPT	SPA	3	7632
	Current Procedural Terminology, 2007	CPT2007	CPT	ENG	3	8847
	CRISP Thesaurus, 2006	CSP2006	CSP	ENG	0	16702
	COSTART, 1995	CST95	CST	ENG	0	3847
	Common Terminology Criteria for Adverse Events, 2003	CTCAEV3	CTCAE	ENG	0	5598
	Diseases Database, 2000	DDB00	DDB	ENG	3	169
	German translation of ICD10, 1995	DMDICD10_1995	ICD10	GER	1	11308

Highlighted rows are excluded from the subset.

Precedence Tab

- Ranks names by types of terms within sources
- Highest ranking name determines the Preferred Name



Full Source Name	Source Abbreviation	Term Type
UMLS Metathesaurus	MTH	PN
Medical Subject Headings, 2007_2007_05_01	MSH2007_2007_05_01	MH
Medical Subject Headings, 2007_2007_05_01	MSH2007_2007_05_01	TQ
Medical Subject Headings, 2007_2007_05_01	MSH2007_2007_05_01	PEP
Medical Subject Headings, 2007_2007_05_01	MSH2007_2007_05_01	PEN
Medical Subject Headings, 2007_2007_05_01	MSH2007_2007_05_01	EP
Medical Subject Headings, 2007_2007_05_01	MSH2007_2007_05_01	EN
Medical Subject Headings, 2007_2007_05_01	MSH2007_2007_05_01	XQ
Medical Subject Headings, 2007_2007_05_01	MSH2007_2007_05_01	PXQ
Medical Subject Headings, 2007_2007_05_01	MSH2007_2007_05_01	NM
RxNorm Vocabulary, 07AA_070503F	RXNORM_07AA_070503F	SCD
RxNorm Vocabulary, 07AA_070503F	RXNORM_07AA_070503F	SBD
RxNorm Vocabulary, 07AA_070503F	RXNORM_07AA_070503F	SY

Cut and paste rows to alter the preferred name

Suppressibility Tab

UMLS Metathesaurus Configuration 2007AB

File Edit Options Reset Done Help

Input Options | Output Options | Source List | Precedence | Suppressibility

This screen contains the Suppressibility Filter, which specifies source/term type combinations to be suppressed. Users can customize the subset by selecting and deselecting source/term type combinations. See Help for more information.

Users may proceed to the other Options tabs in any order by clicking on the tabs across the top of the screen or by clicking "Done" on the File bar and selecting "Begin Subset".

Select One or More Suppressible Term Types

Source	Source Abbreviation	Term Type
Logical Observation Identifier Names and Codes, 219	LNC219	LN
Logical Observation Identifier Names and Codes, 219	LNC219	LO
Logical Observation Identifier Names and Codes, 219	LNC219	LPDN
Logical Observation Identifier Names and Codes, 219	LNC219	LPN
Logical Observation Identifier Names and Codes, 219	LNC219	LS
Logical Observation Identifier Names and Codes, 219	LNC219	LX
Logical Observation Identifier Names and Codes, 219	LNC219	OLX
Logical Observation Identifier Names and Codes, 219	LNC219	OOSN
Logical Observation Identifier Names and Codes, 219	LNC219	OSN
Logical Observation Identifier Names and Codes, 219	LNC219	SN
Logical Observation Identifier Names and Codes, 219	LNC219	SX
Logical Observation Identifier Names and Codes, 219	LNC219	XM
McMaster University Epidemiology Terms, 1992	MCM92	PT
McMaster University Epidemiology Terms, 1992	MCM92	RT
MedlinePlus Health Topics_2004_08_14, 20040814	MEDLINEPLUS_20040814	ET
MedlinePlus Health Topics_2004_08_14, 20040814	MEDLINEPLUS_20040814	PT
MedlinePlus Health Topics_2004_08_14, 20040814	MEDLINEPLUS_20040814	XM
Medical Subject Headings, 2007_2007_05_01	MSH2007_2007_05_01	CE
Medical Subject Headings, 2007_2007_05_01	MSH2007_2007_05_01	DEV

Highlighted source term types will be marked as suppressible

File menu

UMLS Metathesaurus Configuration 2007AB

File Edit Options Reset Done Help

Enable/Disable Filter Ctrl+E
Import Filter.. Ctrl+I
New Configuration.. Ctrl+N
Open Configuration.. Ctrl+O
Save Configuration.. Ctrl+S
Exit Ctrl+Q

Source List Precedence Suppressibility

set data was saved and indicate the location of the source folder that contains the existing subset configuration.

er Options tabs in any order by clicking on the tabs across the top of the screen or by clicking "Done" on the File bar and selecting "Begin Subset".

Input Format Options

Select Input Format

NLM Data File Format Browse..

NLM Data File Format Configuration

Source Folder - Location of Metathesaurus Files

E:\ Browse..

Edit menu

UMLS Metathesaurus Configuration 2007AB

File | Edit | Options | Reset | Done | Help

Undo Suppressibility action	Ctrl+Z
Redo	Ctrl+Y
Increase Font Size	Ctrl+Close Bracket
Decrease Font Size	Ctrl+Open Bracket

ence | Suppressibility

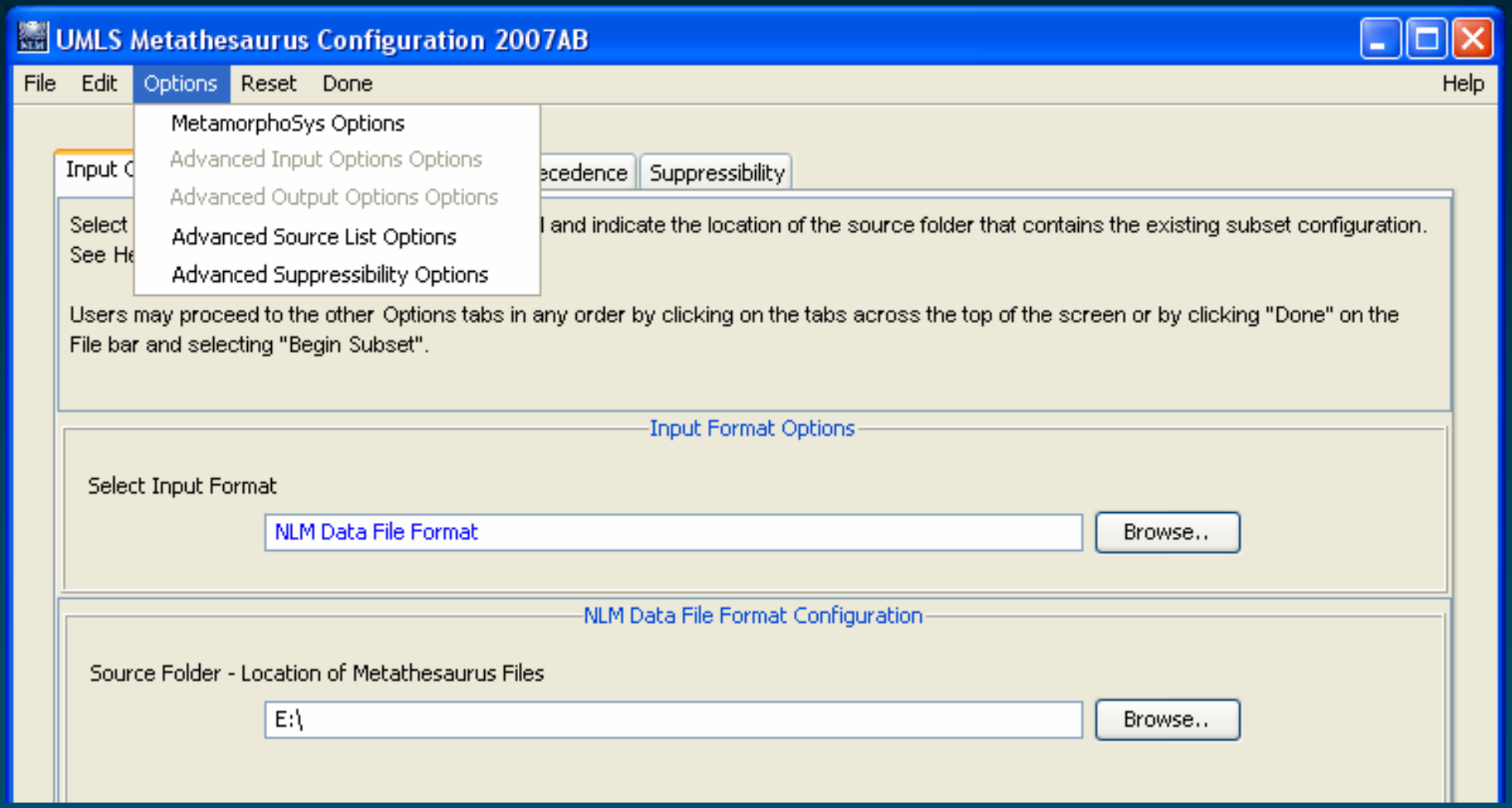
specifies source/term type combinations to be suppressed. Users can customize the combinations. See Help for more information.

Users may proceed to the other Options tabs in any order by clicking on the tabs across the top of the screen or by clicking "Done" on the File bar and selecting "Begin Subset".

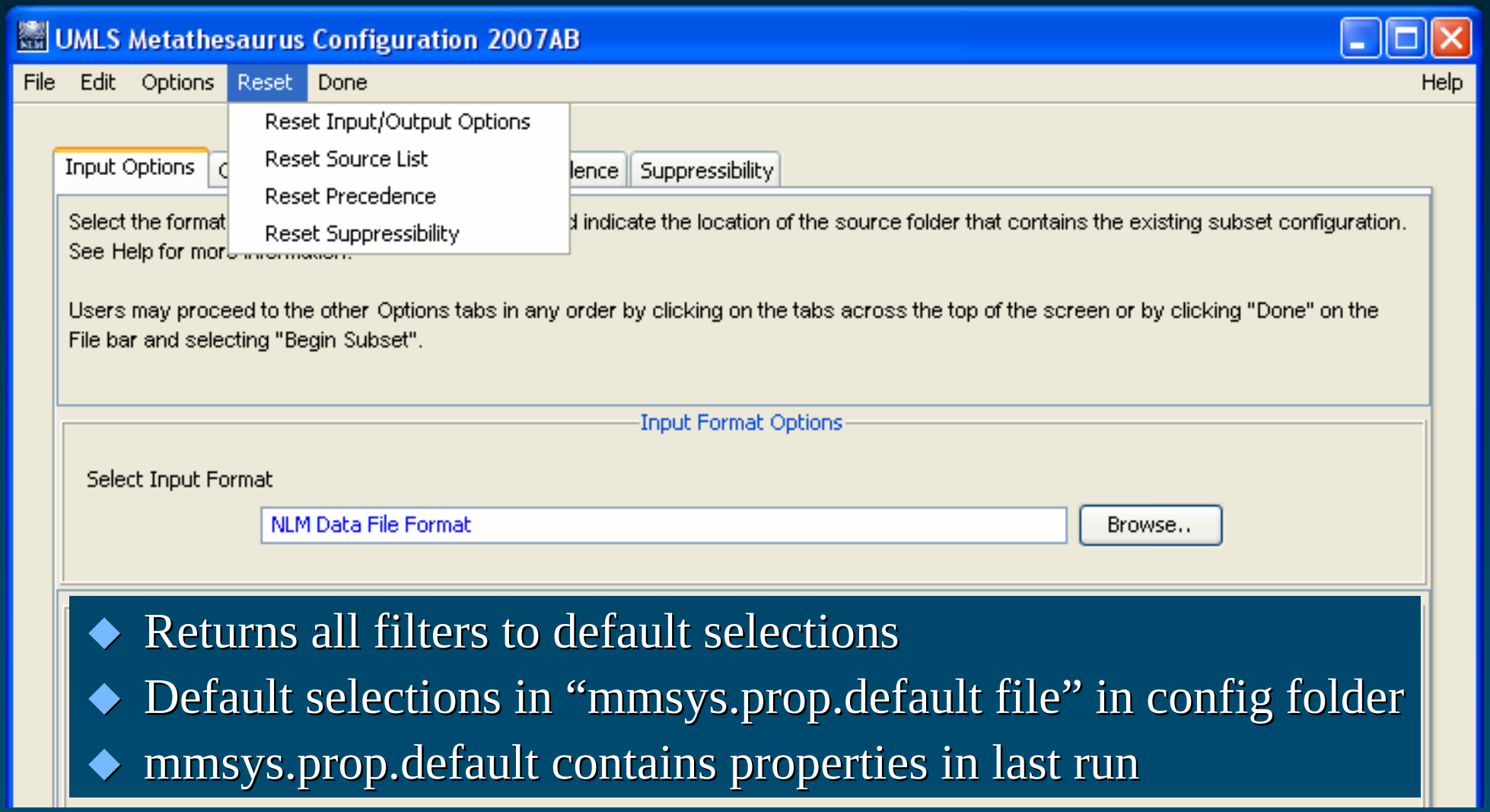
Select One or More Suppressible Term Types

Source	Source Abbreviation	Term Type
Logical Observation Identifier Names and Codes, 219	LNC219	LN
Logical Observation Identifier Names and Codes, 219	LNC219	LO
Logical Observation Identifier Names and Codes, 219	LNC219	LPDN
Logical Observation Identifier Names and Codes, 219	LNC219	LPN
Logical Observation Identifier Names and Codes, 219	LNC219	LS
Logical Observation Identifier Names and Codes, 219	LNC219	LX
Logical Observation Identifier Names and Codes, 219	LNC219	OLX
Logical Observation Identifier Names and Codes, 219	LNC219	OOSN
Logical Observation Identifier Names and Codes, 219	LNC219	OSN
Logical Observation Identifier Names and Codes, 219	LNC219	SN
Logical Observation Identifier Names and Codes, 219	LNC219	SX

Options menu



Reset menu



- Returns all filters to default selections
- Default selections in “mmsys.prop.default file” in config folder
- mmsys.prop.default contains properties in last run

Done – Begin Subset

UMLS Metathesaurus Configuration 2007AB

File Edit Options Reset Done Help

Begin Subset Ctrl+B

Input Options | Output Options | Source List | Precedence | Suppressibility

Select the format in which subset data was saved and indicate the location of the source folder that contains the existing subset configuration. See Help for more information.

Users may proceed to the other Options tabs in any order by clicking on the tabs across the top of the screen or by clicking "Done" on the File bar and selecting "Begin Subset".

Input Format Options

Select Input Format

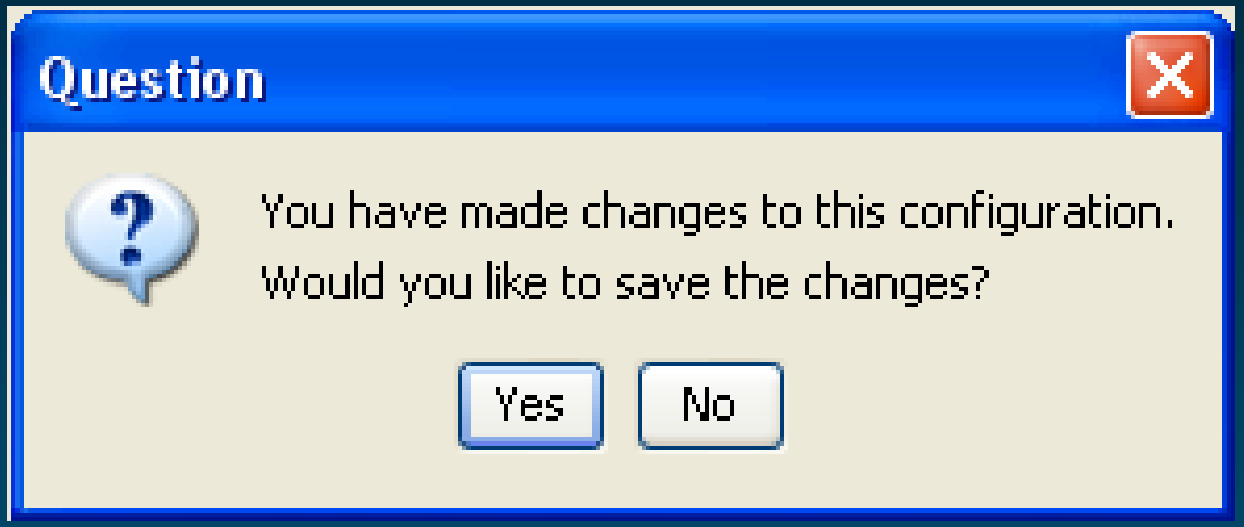
NLM Data File Format [Browse..]

NLM Data File Format Configuration

Source Folder - Location of Metathesaurus Files

E:\ [Browse..]

Save configuration for next installation



How MetamorphoSys Works

- Removes all information from relational files in excluded vocabularies
  - atoms, strings, relationships, attributes, mappings, etc.
- Applies additional options selected by user
  - such as adding source term suppressibility or altering precedence
- Produces a full set of Metathesaurus files
  - relational files with customized data
  - reflecting other user criteria

MetamorphoSys log

MetamorphoSys Subset Log

```
MetamorphoSys Version:...............MMSYS-2007AB-20070621
MetamorphoSys Build Date:............2007_06_21_18_35_25
UMLS Build Date:.....................2007_06_29_11_50_02
Release Version:.....................2007AB
Release Date:........................20070616
Release Description:.................Base Release for Spring 2007
Metathesaurus Source paths:..........E:\
Subsetted Metathesaurus folder:.......D:\2007AB\META
Configuration file used:.............D:\2007AB\ob-mesh.properties
Start at:............................Fri Jul 27 16:28:35 EDT 2007
Initialize CUI List completed:........Fri Jul 27 16:33:00 EDT 2007
Subset Metathesaurus completed:.......Fri Jul 27 17:28:55 EDT 2007
Subset Index Files completed:.........Fri Jul 27 17:36:34 EDT 2007
Subset Release Metadata completed:....Fri Jul 27 17:36:56 EDT 2007
Finished at:.........................Fri Jul 27 17:39:02 EDT 2007
Concepts in source:..................1436586
Concepts in subset:..................278431
Time elapsed:........................01:10:26

Details about MetamorphoSys Input Handler
```

Done | Browse Subset..

MetamorphoSys log

mmsys.log - Notepad

File Edit Format View Help

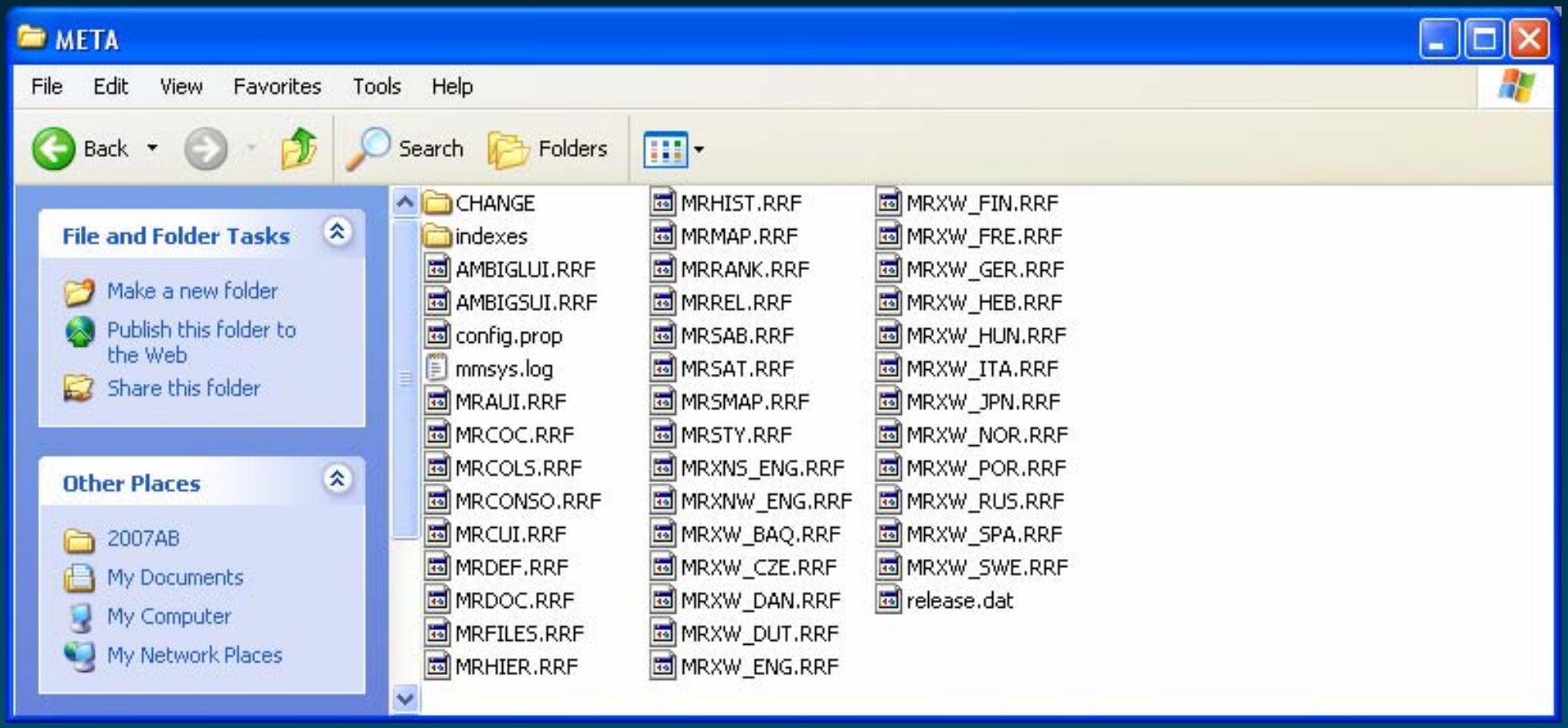
```
MetamorphoSys Version:................MMSYS-2007AB-20070621
MetamorphoSys Build Date:.............2007_06_21_18_35_25
UMLS Build Date:......................2007_06_29_11_50_02
Release Version:......................2007AB
Release Date:.........................20070616
Release Description:..................Base Release for Spring 2007
Metathesaurus Source paths:...........E:\
Subsetted Metathesaurus folder:.......D:\2007AB\META
Configuration file used:..............D:\2007AB\ob-mesh.properties
Start at:.............................Fri Jul 27 16:28:35 EDT 2007
Initialize CUI List completed:........Fri Jul 27 16:33:00 EDT 2007
Subset Metathesaurus completed:.......Fri Jul 27 17:28:55 EDT 2007
Subset Index Files completed:.........Fri Jul 27 17:36:34 EDT 2007
Subset Release Metadata completed:....Fri Jul 27 17:36:56 EDT 2007
Finished at:..........................Fri Jul 27 17:39:02 EDT 2007
Concepts in source:...................1436586
Concepts in subset:...................278431
Time elapsed:.........................01:10:26

Details about MetamorphoSys Input Handler
-----------------------------
MetamorphoSys Input: NLM Data File Format

Details about MetamorphoSys Output Handler
------------------------------
Metathesaurus Output: Rich Release Format
Long fields were not truncated.
Source Abbreviations were written out with a versionless (root)
representation.
Fields containing UTF-8 characters were not removed.
UTF-8 BOM characters were not added to output files.
Md5s were not calculated.
Concepts containing only MTH atoms were not removed.


Excluded Sources
  AI/RHEUM, 1993
```

Output directory contents



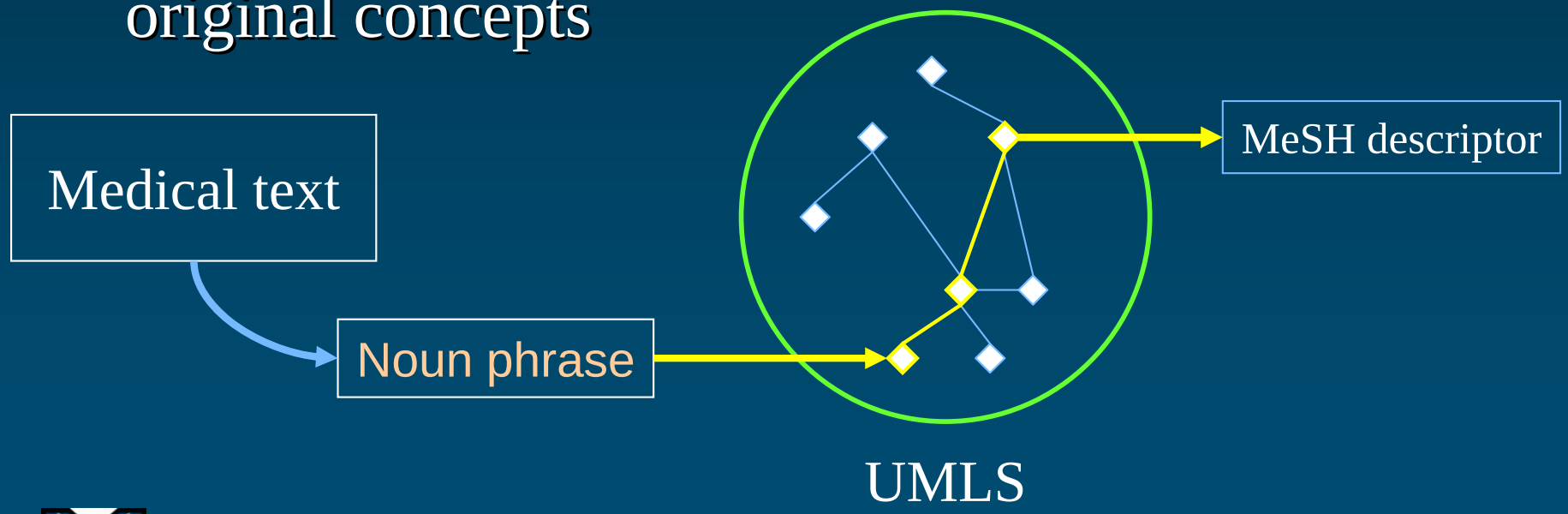
Part II
How to use the UMLS?

(4) A UMLS-based algorithm

Indexing Initiative

[Aronson & al., *AMIA*, 2000]

- For noun phrases extracted from medical texts, map to UMLS concepts
- Then, select from the MeSH vocabulary the concepts that are the most closely related to the original concepts



Restrict to MeSH

[Bodenreider & al., *AMIA*, 1998]

- Based on the principle of semantic locality
- Use different components of the UMLS
- 4 techniques of increasing aggressiveness
  - Use Synonymy — MRCONSO
  - Use Associated expressions (ATXs) — MRATX + MRREL
  - Explore the Ancestors — MRREL + SN
  - Explore the Other related concepts — MRREL + SN

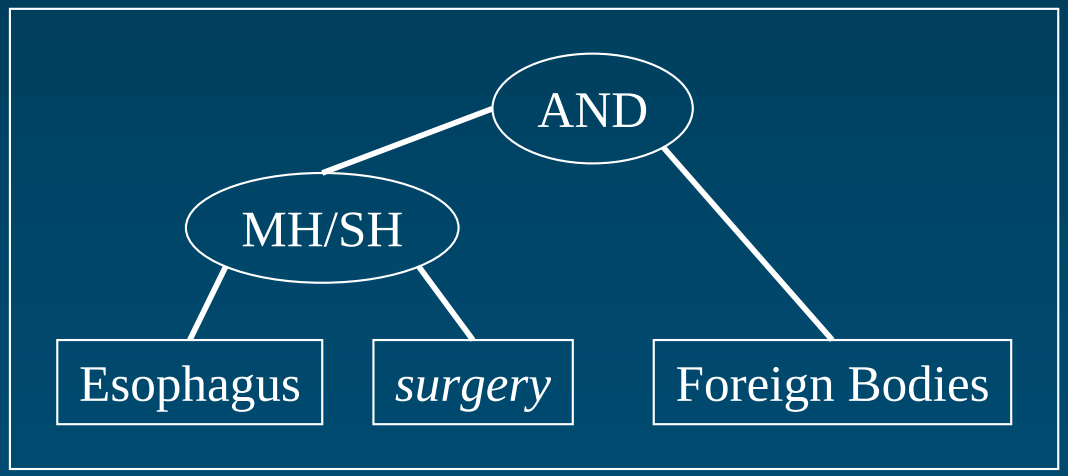
Restrict to MeSH Synonymy

- Term mapped to Source concept
- For this concept, is there a synonym term that comes from MeSH? (MRCONSO)

Restrict to MeSH Assoc. expressions

- If not,
- Is there an associated expression (ATX) that describes this concept using a combination of MeSH descriptors? (MRATX/MRMAP + MRREL)

Endoscopic removal of intraluminal foreign body from oesophagus without incision



Restrict to MeSH Ancestors

- If not, let us build the graph of the ancestors of this concept
  - using parents and broader concepts (MRREL)
  - all the way to the top
  - excluding ancestors whose semantic types are not compatible with those of the source concept (MRSTY)
- From the graph, select the concepts that come from MeSH (MRCONSO)
- Remove those that are ancestors of another concept coming from MeSH

Restrict to MeSH Other related concepts

- If not, explore the other related concepts (MRREL) whose semantic types are compatible with those of the source concept (MRSTY)
- From those, select the concepts that come from MeSH (MRCONSO)

Restrict to MeSH Example

Vein of neck, NOS

There is a MeSH term in the synonyms of SC

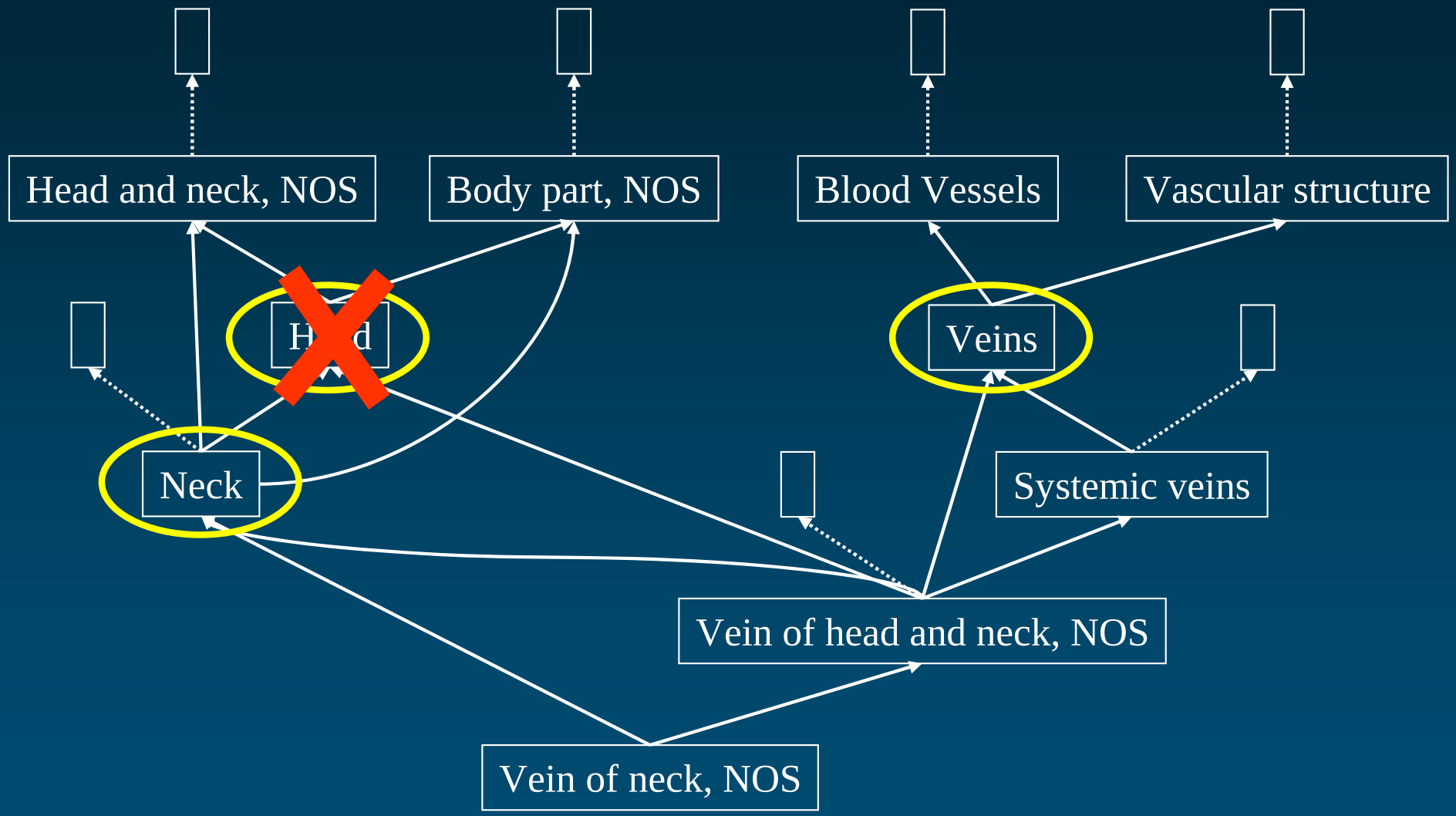
SC is described by a combination of MeSH terms (ATX)

The ancestors of SC contain MeSH terms

MeSH terms from non-hierarchically related concepts

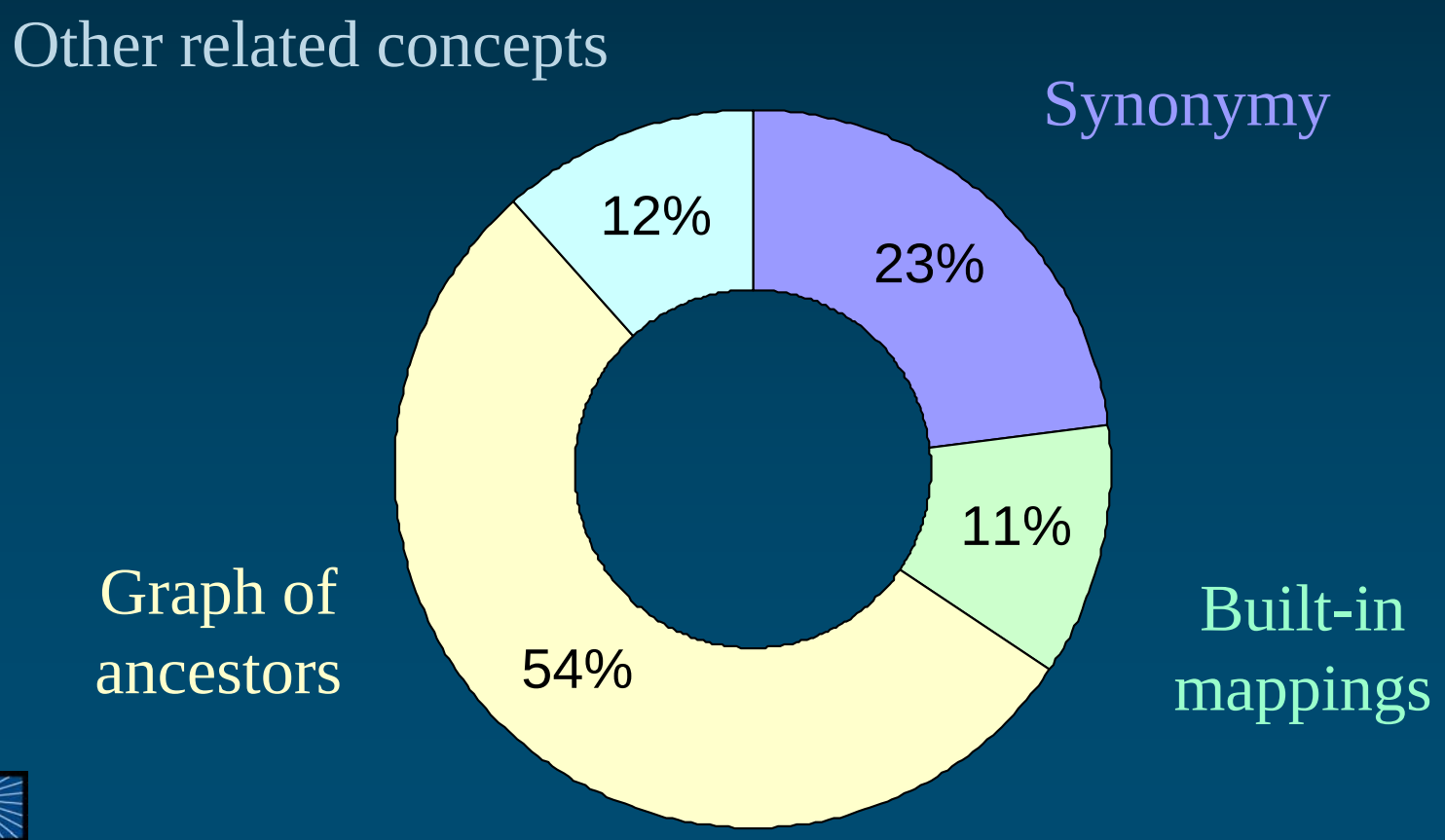
Vein + Neck

Restrict to MeSH Example



Restrict to MeSH Quantitative results

- 86% of UMLS concepts mapped to MeSH (2007)



Restrict to MeSH Qualitative results

- Qualitative evaluation
  - 1,036 concepts extracted from 200 MEDLINE citations
  - manual review of every mapping or failure
- 61% Relevant
  - Subtotal Gastrectomy → Gastrectomy
  - Encephalopathy, NOS → Brain Diseases
- 28% More or less relevant
  - Vitamin A measurement → Laboratory Procedure
  - Swelling, NOS → Symptoms
- 11% Non relevant

Part II
How to use the UMLS?

(5) Benefits and Limitations

Benefits

UMLS compared to individual vocabularies

- Broader scope
- Extended coverage
- Finer granularity
- Unique identifier
- Synonymous terms clustered into concepts
- Additional synonyms
- Additional hierarchical relationships
- Semantic categorization

Direct benefits

- Concept categorization
- Information retrieval
  - Synonyms
  - Cross-language features
- Information extraction
  - MetaMap
  - Normalization
- Information visualization
  - Knowledge Source Server
  - Semantic Navigator
  - RRF browser

UMLS as an enabling resource

- Examples
  - Mapping across vocabularies
  - Semantics of statistical associations
  - Redundancy in hierarchical relations

Limitations

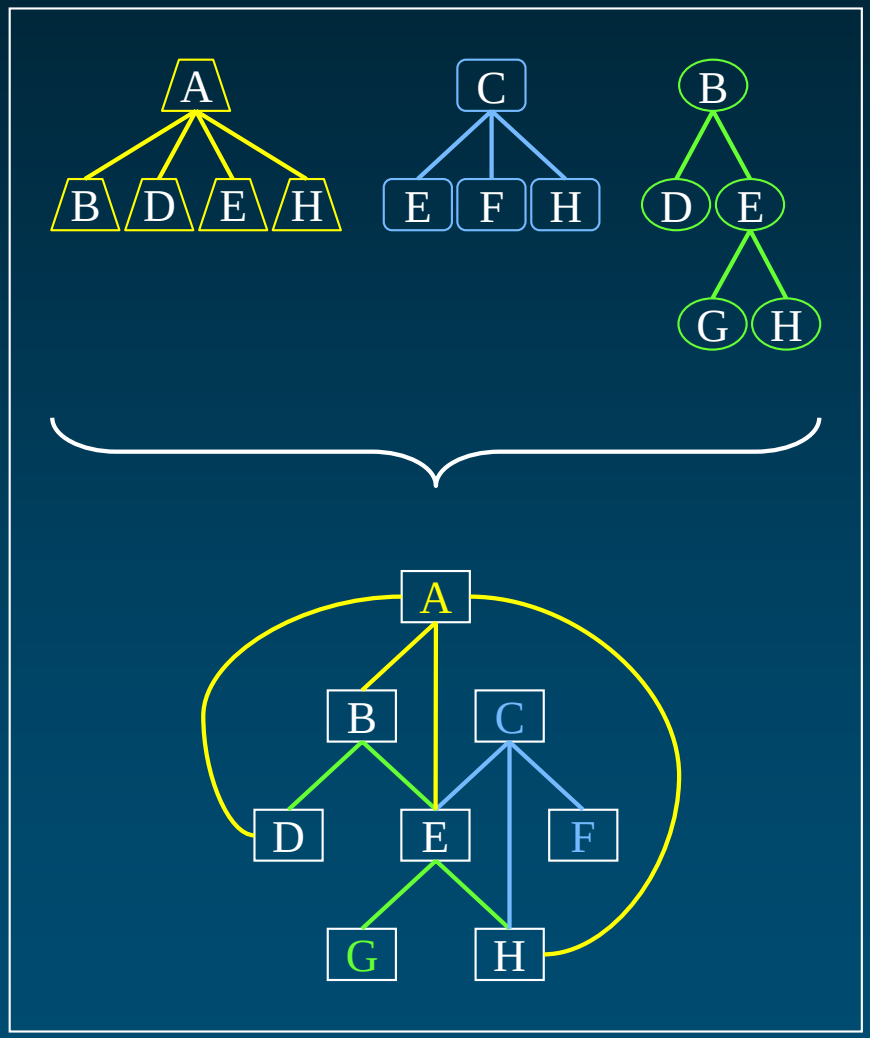
Limitations

[Cimino, *JAMIA*, 1998]

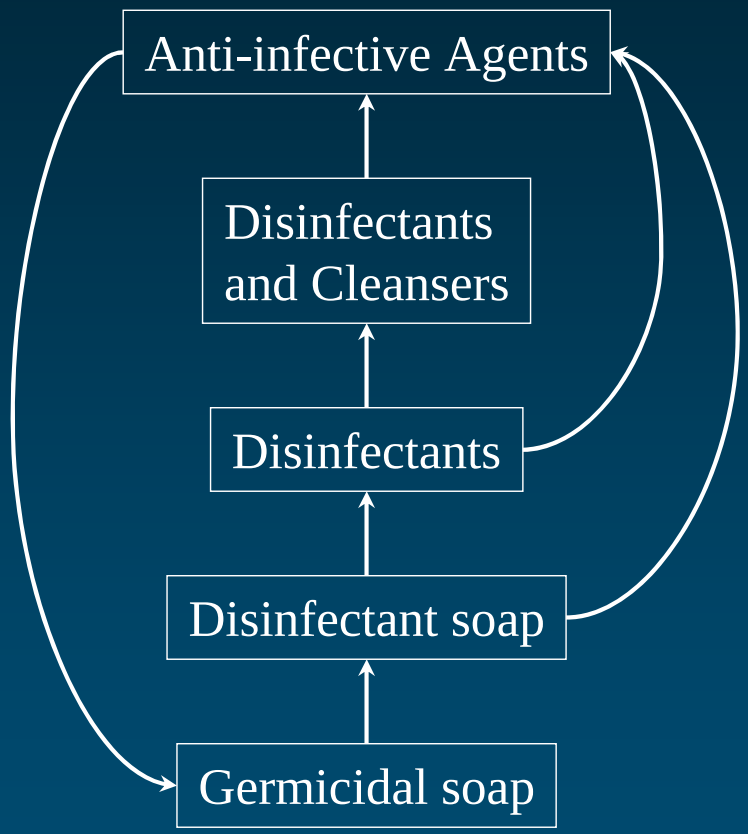
- Structural inconsistency
  - Cycles in the graph of hierarchical relations
- Semantic inconsistency
  - Between Metathesaurus and Semantic Network
- Underspecified relationships
- Missing relations
  - Synonymy
  - Hierarchical relations (missing or underspecified)

Structural inconsistency From trees to graph

- Multiple tree structures combined into a graph structure
- Expected: Directed acyclic graph (DAG)

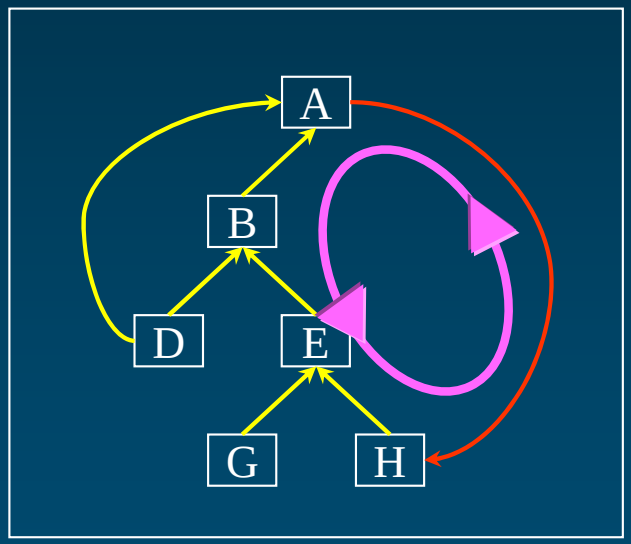


Structural inconsistency Cycles in the UMLS graph



Structural inconsistency Issues

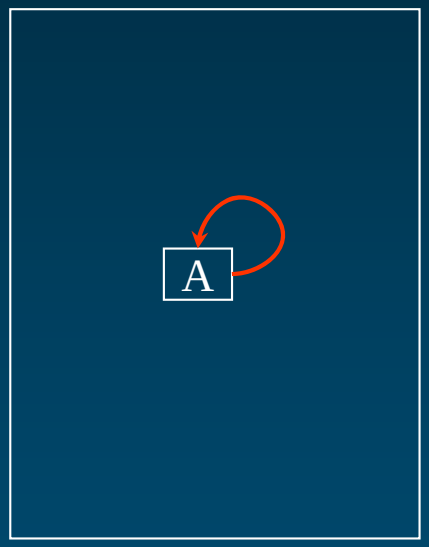
- Theoretical
  - Violate the antisymmetry property of partial ordering relations
- Practical
  - Loops in graph traversal
  - Impossible to perform transitive reduction



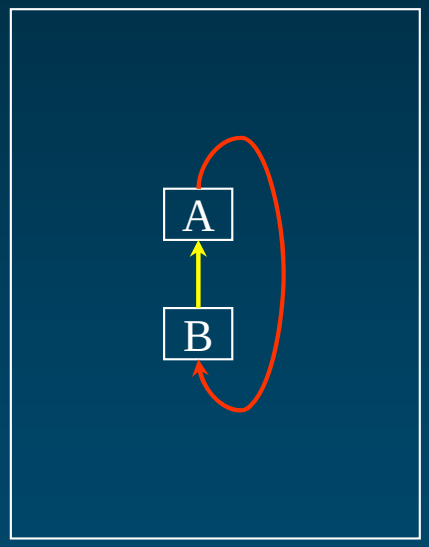
[Bodenreider, AMIA 2001]

Acyclicity

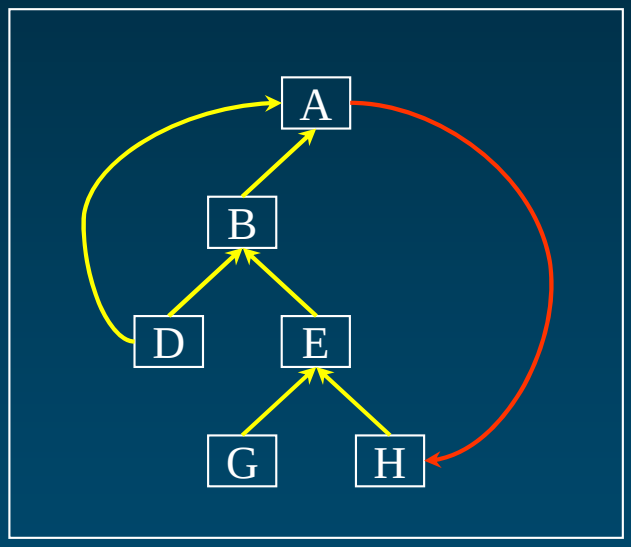
"back edge" from a child concept to a parent concept



Reflexive
13,000

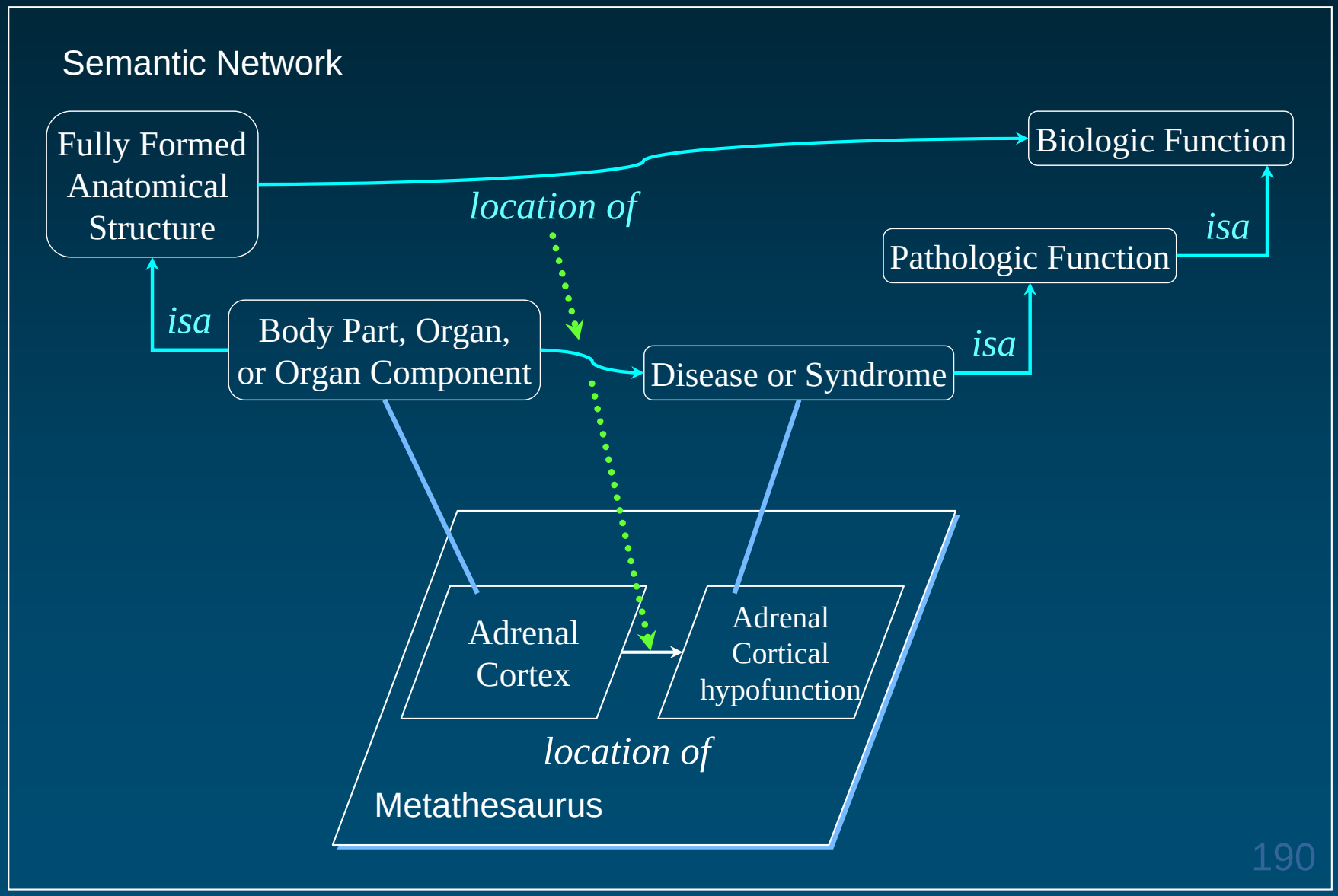


Direct
1800



Indirect
120

Semantic inconsistency A two-level structure

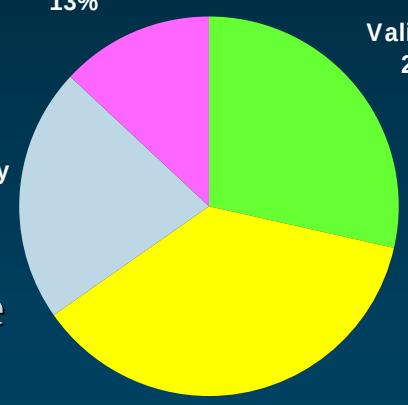


Semantic inconsistency A limited study

- 6894 interconcept relationships
  - among the 3764 concepts in the semantic neighborhood of "Heart"

ICR and SNR not compatible

ICR = SNR or ICR descendant of SNR



ICR not specified and SNR compatible and multiple

ICR not specified and SNR compatible and unique

[McCray A.T, Bodenreider O. A conceptual framework for the biomedical domain. In: Green R, Bean CA, Myaeng SH, editors. *The semantics of relationships: an interdisciplinary perspective*. Boston: Kluwer Academic Publishers; 2002. p. 181-198.]

Semantic inconsistency Issues

- The UMLS integrates what terminologies represent
- Hierarchies in source vocabularies
  - Often task-driven rather than based on principles
  - Usually suitable for information retrieval
  - Not necessarily suitable for reasoning
- No automatic correction possible
  - Wrong categorization
  - Wrong inter-concept relationship
  - [Wrong semantic network relationship]

Underspecified relationships

- Relationship “attribute” not always present
- Relations used to create hierarchies vs. hierachical relations

Environment and Public Health [G03]

Public Health [G03.850]

Accidents [G03.850.110]

Accident Prevention [G03.850.110.060] +

Accidental Falls [G03.850.110.085]

Accidents, Aviation [G03.850.110.185]

Accidents, Home [G03.850.110.205]

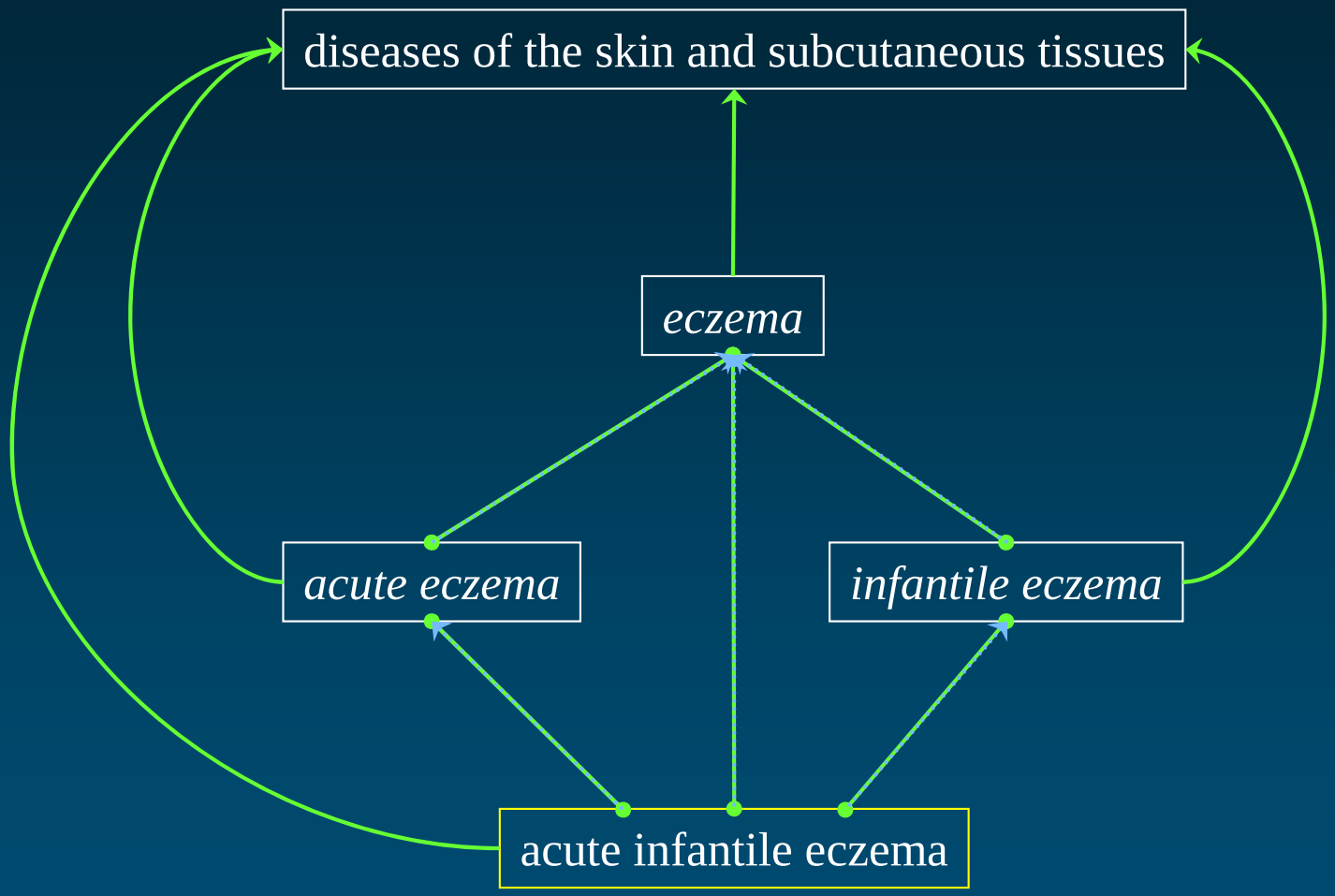
Accidents, Occupational [G03.850.110.250] +

Accidents, Radiation [G03.850.110.285]

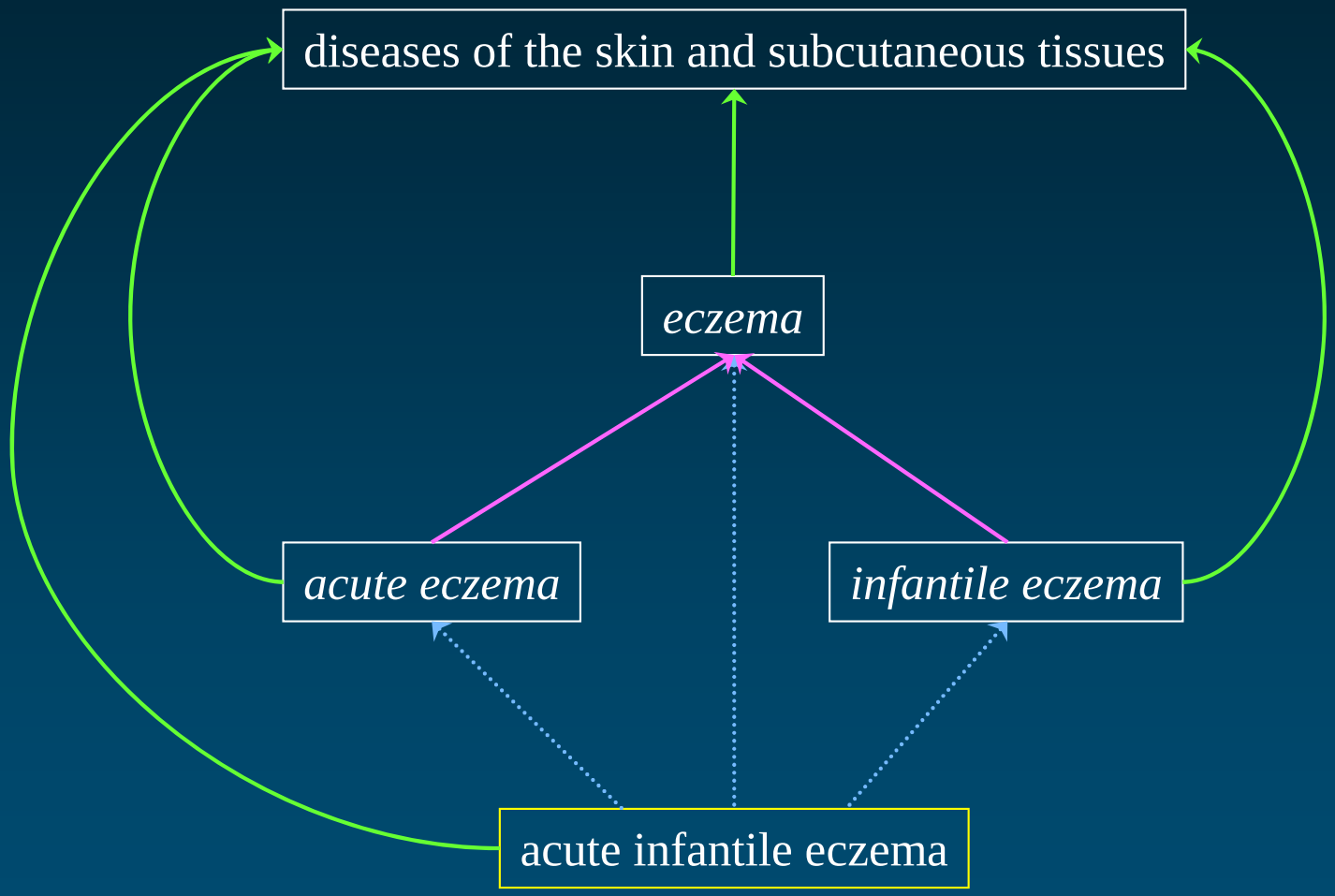
Accidents, Traffic [G03.850.110.320]

Drowning [G03.850.110.500] +

Missing relations Example



Missing relations Example



Missing relations A limited study

- 28,851 pairs of terms
  - Original SNOMED term
  - Demodified term (found in UMLS)
- Corresponding relationship in the Metathesaurus
  - Hierarchical in 50% of the cases
  - « Sibling » in 25% of the cases
  - Missing in 25% of the cases

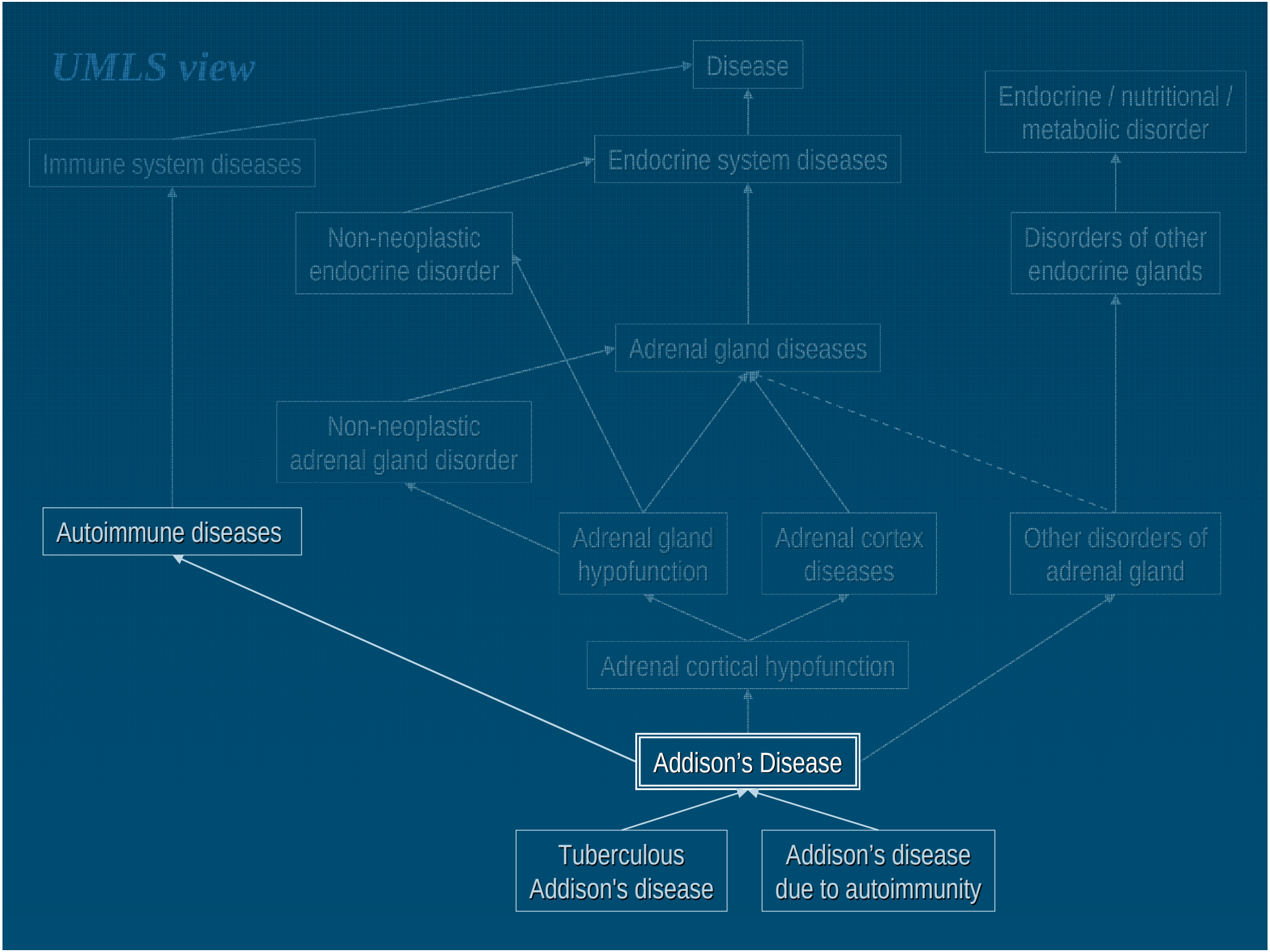
[Bodenreider & al., *TIA*, 2001]

Compensation mechanisms

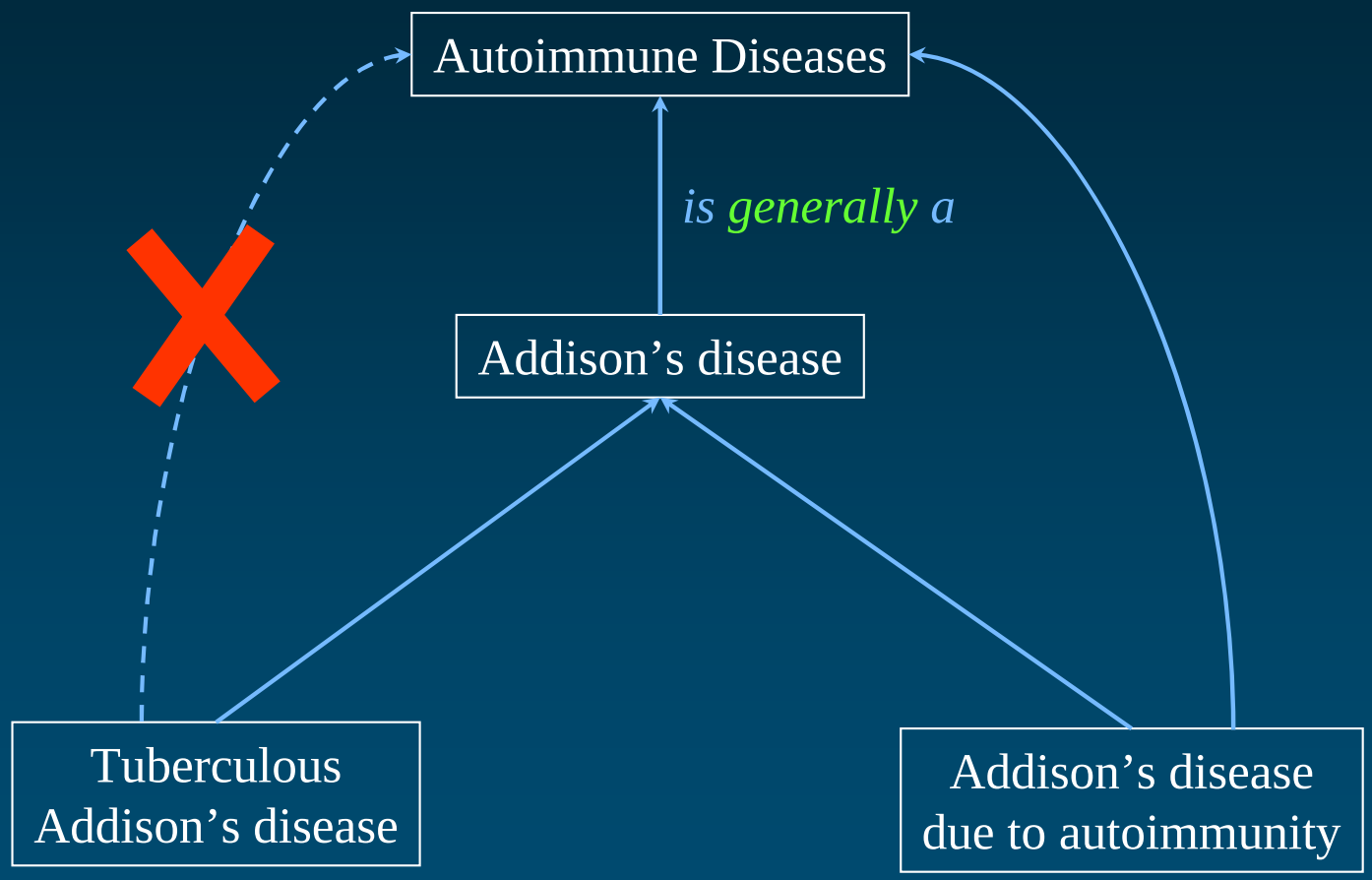
- Examples
  - Removing cycles from hierarchical relations
    - Using redundancy (number of sources asserting the relation)
    - Using terminological knowledge (e.g., NEC)
  - Lexically-suggested hyponymic relations
    - Properties of adjectival modification

More limitations

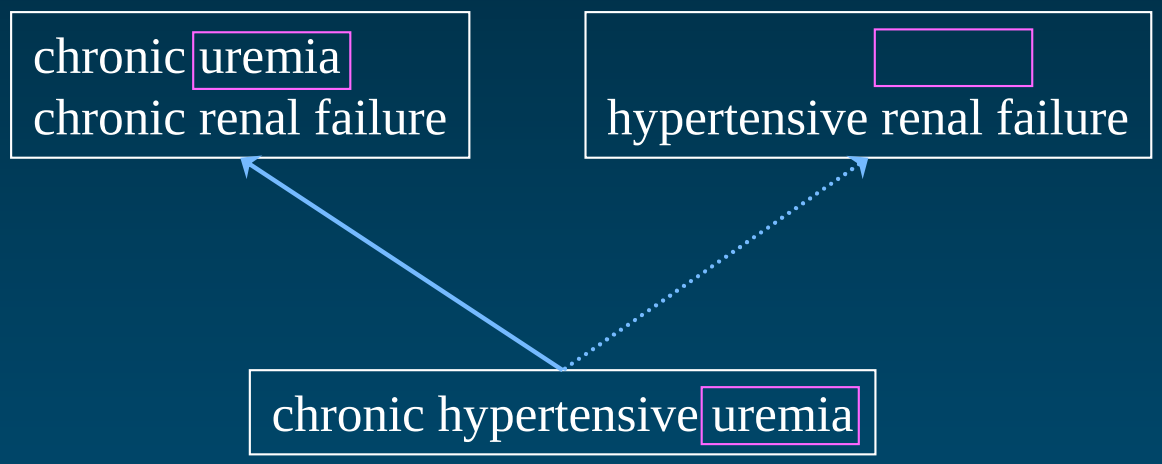
- Semantics of hierarchical relations
- Some missing / wrong relations are hard to detect
- Some relations are present but hard to find


UMLS view
Disease
Endocrine / nutritional / metabolic disorder
Immune system diseases
Endocrine system diseases
Non-neoplastic endocrine disorder
Disorders of other endocrine glands
Adrenal gland diseases
Non-neoplastic adrenal gland disorder
Autoimmune diseases
Adrenal gland hypofunction
Adrenal cortex diseases
Other disorders of adrenal gland
Adrenal cortical hypofunction
Addison's Disease
Tuberculous Addison's disease
Addison's disease due to autoimmunity

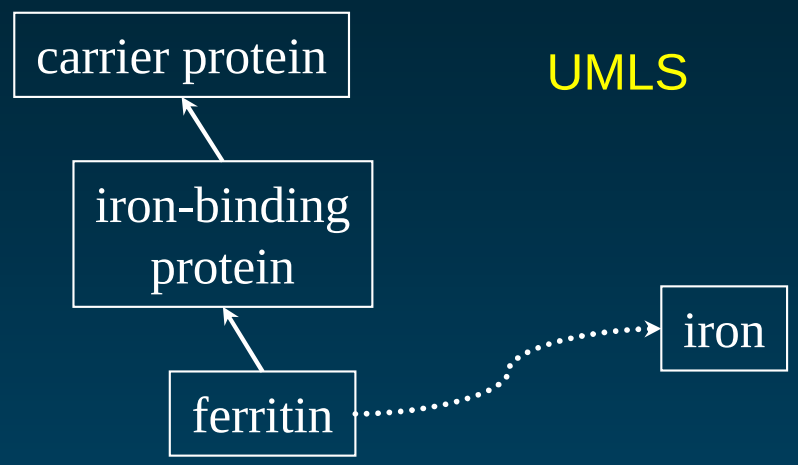
Semantics of hierarchical links



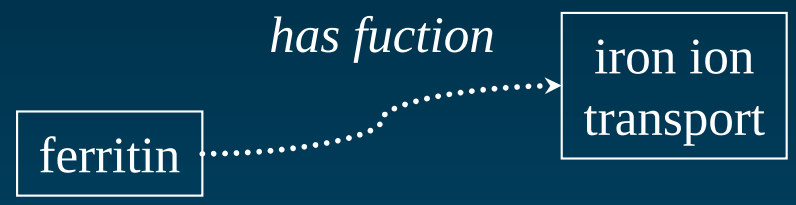
Relations Missing and difficult to detect



Relations Existing but difficult to find



UMLS



Gene Ontology

ferritin *isa* iron transporter

reified “transport” relationship

ferritin *transports* iron

“transport” relationship

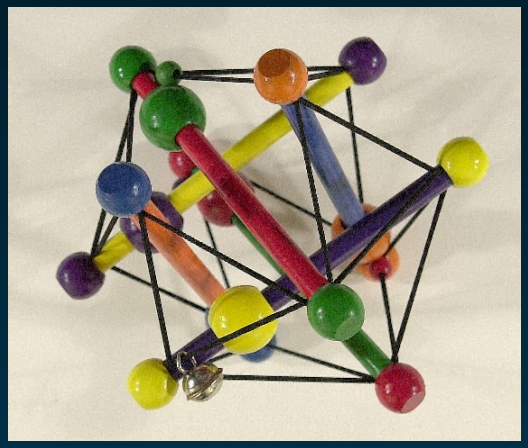
How to address these limitations?

- Description logics
- Natural Language Processing
  (semantic interpretation of the terms)
- Comparing knowledge sources
  (alignment, inference)

Summary

UMLS Summary

- UMLS = 3 Knowledge Sources
  - Metathesaurus
  - Semantic Network
  - SPECIALIST Lexicon and Lexical Tools
- MetamorphoSys
  - installs
  - customizes
- UMLSKS
  - remote access
  - resources and documentation



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

Bibliography

References: UMLS home page

- UMLS home page
  - http:// www.nlm.nih.gov/research/umls/
- UMLS documentation
  - Formerly know as the “Green Book”
  - Now online documentation
  - http://www.nlm.nih.gov/research/umls/UMLSDOC.HTML

References

- Short presentation
  - Bodenreider, O. (2004) The Unified Medical Language System (UMLS): integrating biomedical terminology. *Nucleic Acids Res*, 32(Database issue), D267-70.
- 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.

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.
- Comprehensive bibliography 1986-96

  http://www.nlm.nih.gov/pubs/cbm/umlscbm.html

Documentation and Support

UMLS documentation and support

- UMLS homepage
  - links to various UMLS resources
  - http://www.nlm.nih.gov/research/umls/
- UMLSKS homepage
  - links to the User's and Developer's guides
  - http://umlsks.nlm.nih.gov/
- UMLS mailing list
  - UMLSUSERS-L@LIST.NIH.GOV
- Email address for support
  - custserv@nlm.nih.gov

Appendix

UMLS files in
Rich Release Format

MRCONSO (sample rows 1..5)

(2004AB)

	1	2	3	4	5	6	7	8	9	10	11
	CUI	LAT	TS	LUI	STT	SUI	ISPREF	AUI	SAUI	SCUI	SDUI
1	C0001403	ENG	P	L0001403	PF	S0354372	Y	A4367951			
2	C0001403	ENG	P	L0001403	PF	S0354372	N	A2922421	485624014	363732003	
3	C0001403	ENG	P	L0001403	VC	S0010794	Y	A0019740		M0000346	D000224
4	C0001403	ENG	S	L0494851	PF	S2164152	N	A2018589			
5	C0001403	FRE	P	L3246333	PF	S3773545	Y	A3996251			D000224

	12	13	14	15	16	17	18
	SAB	TTY	CODE	STR	SRL	SUPPRESS	CVF
1	MTH	PN	NOCODE	Addison's disease	0	N	
2	SNOMEDCT	PT	363732003	Addison's disease	4	N	
3	MSH	MH	D000224	Addison’s Disease	0	N	
4	MDR	LT	10052381	Primary adrenal insufficiency	3	N	
5	MSHFRE	MH	D000224	Addison, maladie	3	N	

MRCONSO (sample rows 6..10)

(2004AB)

	1	2	3	4	5	6	7	8	9	10	11
	CUI	LAT	TS	LUI	STT	SUI	ISPREF	AUI	SAUI	SCUI	SDUI
6	C0001403	FRE	S	L1272481	PF	S1514427	Y	A1464383			
7	C0001403	GER	P	L1229627	PF	S1471573	Y	A4030156			D000224
8	C0001403	GER	S	L1239271	PF	S1481217	Y	A4034094			D000224
9	C0001403	JPN	P	L3437833	PF	S3965327	Y	A4264008			D000224
10	C0001403	JPN	S	L3465347	PF	S3992841	Y	A4291522			D000224

	12	13	14	15	16	17	18
	SAB	TTY	CODE	STR	SRL	SUPPRESS	CVF
6	WHOFRE	IT	0410	MALADIE D'ADDISON	2	N	
7	MSHGER	MH	D000224	Addison-Krankheit	3	N	
8	MSHGER	SY	D000224	Bronzehautkrankheit	3	N	
9	MSHJPN	MH	D000224	Addison病	3	N	
10	MSHPJN	SY	D000224	副腎性黒皮症	3	N	

MRCONSO (sample rows 11-13) (2004AB)

	1	2	3	4	5	6	7	8	9	10	11
	CUI	LAT	TS	LUI	STT	SUI	ISPREF	AUI	SAUI	SCUI	SDUI
11	C0001403	POR	P	L3302998	PF	S3831123	N	A6382080			
12	C0001403	RUS	P	L3336992	PF	S3864473	Y	A4157629			
13	C0001403	SPA	P	L1226877	PF	S1468823	Y	A1419475			

	12	13	14	15	16	17	18
	SAB	TTY	CODE	STR	SRL	SUPPRESS	CVF
11	MDRPOR	LT	1001130	Doença de Addison	3	N	
12	MSHRUS	MH	D000224	АДДИСОНОВА БОЛЕЗНЬ	3	N	
13	WHOSPA	IT	0410	ADDISON, ENFERMEDAD	3	N	

MRHIER (sample rows)

(2004AB)

	1	2	3	4	5	6
	CUI	AUI	CXN	PAUI	SAB	RELA
1	C0001403	A0019740	1	A0020270	MSH	
2	C0001403	A0019740	2	A0028022	MSH	
3	C0001403	A0019743	3	A1988358	PSY	member_of_cluster
4	C0001403	A2922421	1	A3307650	SNOMEDCT	isa
5	C0001403	A2922421	2	A3307650	SNOMEDCT	isa

	7	8	9
	PTR	HCD	CVF
1	A0434168.A2367943.A2366890.A0135391.A0054194.A0020267.A0020270	C19.053.264.263	
2	A0434168.A2367943.A2366890.A0135391.A0072566.A0028022	C20.111.163	
3	A0449751.A1988279.A1988358		
4	A3684559.A3886745.A2880798.A3398606.A3399335.A3398961.A2872359.A2872360.A3307650		
5	A3684559.A3886745.A2880798.A3398606.A3399335.A3398961.A2872359.A2933400.A2989549.A3307650		

MRREL (sample rows)

(2004AB)

	1	2	3	4	5		6
	CUI1	AUI1	STYPE1	REL	CUI2	AUI2	STYPE2
1	C0001403		CUI	RB	C0001621		CUI
2	C0001403	A0019738	AUI	SY	C0001403	A0049628	AUI
3	C0001403	A2922421	SCUI	CHD	C0085859	A2977940	SCUI
4	C0001403	A6326321	SCUI	RO	C0688490	A6339383	SCUI
5	C0001403	A0019743	AUI	PAR	C0935495	A1988358	AUI

	7	8	9	10	11	12	13	14	15
	RELA	RUI	SRUI	SAB	SL	RG	DIR	SUPPRESS	CVF
1		R02837989		MTH			N	N	
2		R18849683		MSH	MSH			N	
3	isa	R19859511	1658795027	SNOMEDCT	SNOMEDCT	0	Y	N	
4	may_treat	R27600039		NDFRT	NDFRT			N	
5	has_member	R08110401		PSY	PSY			N	

MRDEF

(2004AB)

CUI	AUI	ATUI	SATUI	SAB	DEF	SUPPRESS	CVF
C0001403	A0019740	AT15061584		MSH	A disease characterized by hypotension, weight loss, anorexia, weakness, and sometimes a bronze-like melanotic hyperpigmentation of the skin. It is due to tuberculosis- or autoimmune-induced disease (hypofunction) of the adrenal glands that results in deficiency of aldosterone and cortisol. In the absence of replacement therapy, it is usually fatal.	N	

MRSAT (sample rows)

(2004AB)

	1	2	3	4	5	
	CUI	LUI	SUI	METAUI	STYPE	CODE
1	C0001403	L0001403	S0010792	A0019738	AUI	D000224
2	C0001403	L0001403	S0010794	A6326321	SCUI	C712
3	C0001403	L0001403	S0354372	A2922421	SAUI	363732003
4	C0001403			R15742591	SRUI	
5	C0001403				CUI	

	7	8	9	10	11	14	15
	ATUI	SATUI	ATN	SAB	ATV	SUPPRESS	CVF
1	AT15321482		DID	MSH	D000224	N	
2	AT33411754		MESH_UI	NDFRT	D000224	N	
3	AT24166602		DESCRIPTION STATUS	SNOMEDCT	0	N	
4	AT27438950		REFINABILITY	SNOMEDCT	0	N	
5	AT02925340		ST	MTH	R	N	

MRSTY

(2004AB)

CUI	TUI	STN	STY	ATUI	CVF
C0001403	T047	B2.2.1.2.1	Disease or Syndrome	AT17683850	

MRHIST (sample rows)

(2004AB)

	1	2	3	4	5
	CUI	SOURCEUI	SAB	SVER	CHANGETYPE
1	C0001403	1198962018	SNOMEDCT	20020731	0
2	C0001403	1212124016	SNOMEDCT	20020731	0
3	C0001403	1490869013	SNOMEDCT	20030131	0
4	C0001403	363732003	SNOMEDCT	20020129	0
5	C0001403	373662000	SNOMEDCT	20020731	0

	6	7	8	9
	CHANGEKEY	CHANGEVAL	REASON	CVF
1	DESCRIPTIONSTATUS	0		
2	DESCRIPTIONSTATUS	0		
3	DESCRIPTIONSTATUS	0		
4	CONCEPTSTATUS	0		
5	CONCEPTSTATUS	0		