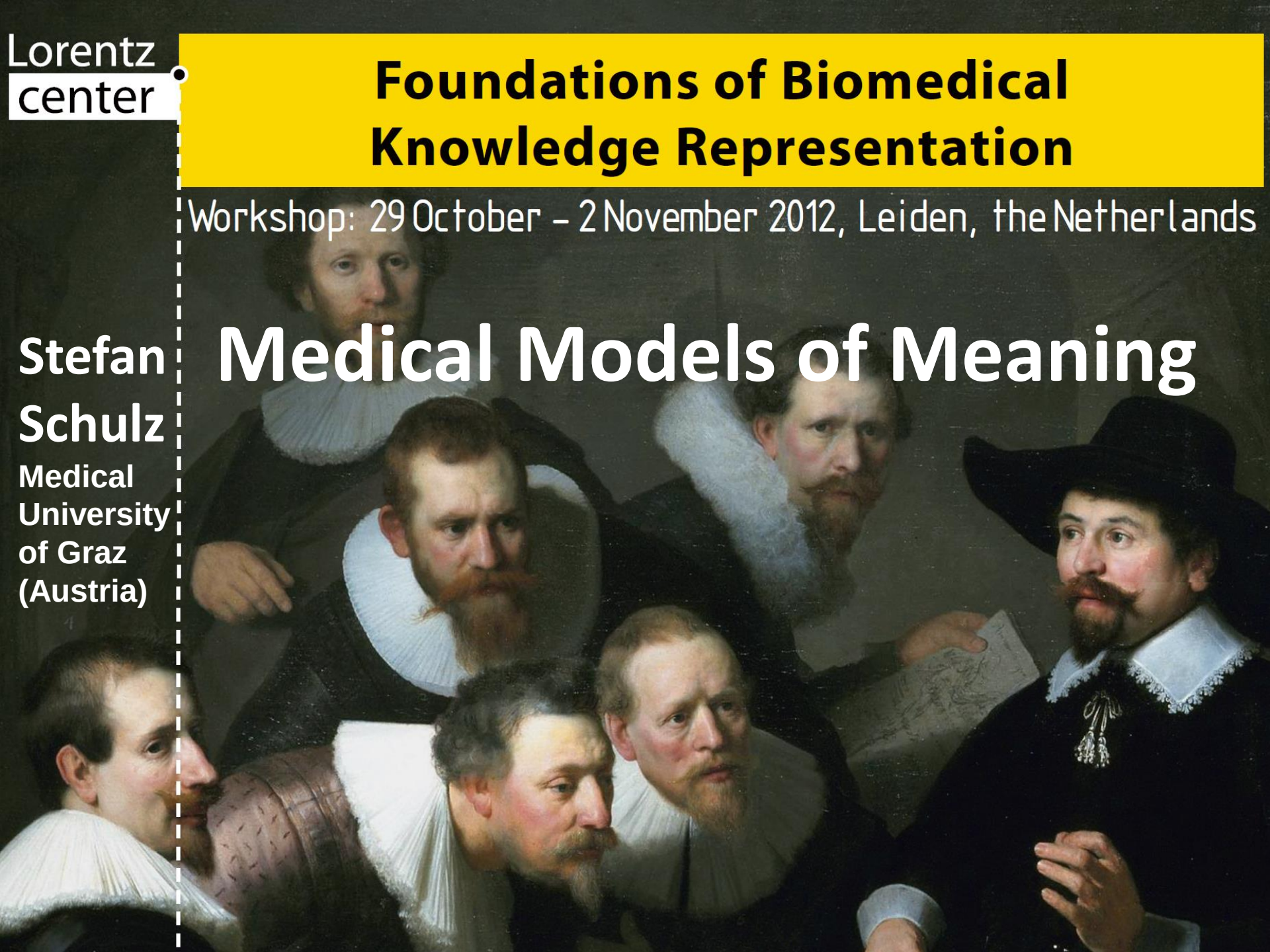
Lorentz center

Foundations of Biomedical Knowledge Representation

Workshop: 29 October – 2 November 2012, Leiden, the Netherlands

Medical Models of Meaning

Stefan Schulz
Medical University of Graz (Austria)

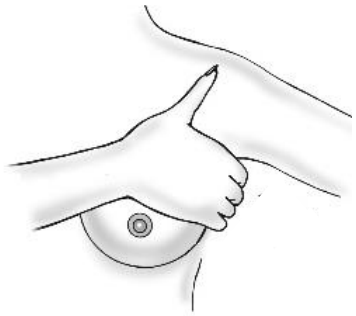


Objectives

- To demonstrate facets of "meaning" in health care and biomedical science
- To outline the scope of "models of meaning" (terminologies, thesauri, classifications, **ontologies**, information models) in representing medical meaning
- To present important biomedical semantic resources, standards and representational formalisms
- To stimulate a discussion about
  - prospects and limitations of "models and meaning" in representing (i) clinical facts and (ii) biomedical knowledge
  - the interface between models of meaning and medical knowledge representation

Clinical example (I)

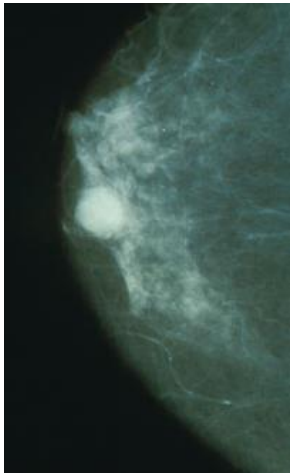
- Female, 45 years old,
  detects small lump in left breast
  (self exam).



- GP:
  - "cherry-sized painless lump in upper left quadrant of left breast"
  - no previous history of neoplasms
  - referral to specialist:
    "breast ca?"

Clinical example (II)

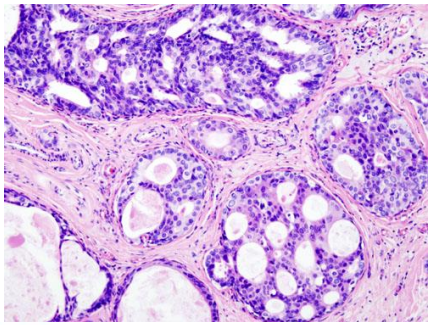
- Specialist:
  - history of "breast nodes": fibroadenoma?
  - family history of breast cancer (mother, diagnosed at 51, total hemimastectomy, brain metastases, death with 59)
  - palpation: painless lump (1cm)

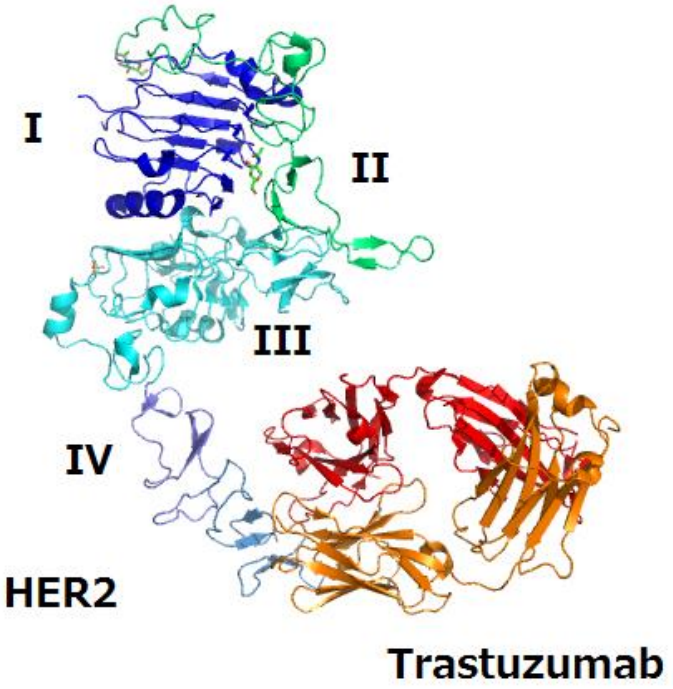


- Routine lab: no abnormalities
- Mammogram: Suggestive of Malignancy
- scheduled for Lumpectomy

Therapy + follow up (III)

- Surgical removal of lump (1.3cm) from left breast
  - Histology: invasive ductal carcinoma, HER2+
  - ICD: C50.4
  - ICD-O: M8500
  - TNM: T1N0M0 (0.7 cm)
- Plan
  - chemotherapy: monoclonal antibody trastuzumab (Herceptin) , 1y
  - echocardiography screening
- Information of patient
  - 5y Survival rate: Stage I: 88%
  - known drug side effect: heart disease





Facts and knowledge (I)

Literature search: monoclonal antibodies and cancer therapy

PubMed.gov
US National Library of Medicine
National Institutes of Health

PubMed | ("Antibodies, Monoclonal"[Mesh]) AND "Neoplasms"[Mesh] | Search

RSS Save search Advanced Help

Show additional filters

Clear all

Text availability
Abstract available
Free full text available
Full text available

Publication dates
5 years
10 years
Custom range...

Species
Humans
Other Animals

Article types clear
Clinical Trial
Meta-Analysis
Practice Guideline
Randomized Controlled Trial
✓ Review
Systematic Reviews
more ...

Display Settings: Summary, 20 per page, Sorted by Recently Added **Send to:**

Results: 2021 to 2040 of 5979

<< First < Prev Page 102 of 299 Next > Last >>

Filters activated: Review, English Clear all

2021. Trastuzumab as the lead monoclonal antibody in advanced breast cancer: choosing which patient and when.
Olver IN.
Future Oncol. 2008 Feb;4(1):125-31. **Review**.
PMID: 18241007 [PubMed - indexed for MEDLINE]
Related citations

2022. Predicting outcome to VEGF-targeted therapy in metastatic clear-cell renal cell carcinoma: data from recent studies.
Golshayan AR, Brick AJ, Choueiri TK.
Future Oncol. 2008 Feb;4(1):85-92. **Review**.
PMID: 18241003 [PubMed - indexed for MEDLINE]
Related citations

2023. Rituximab: beyond simple B cell depletion.
Kessel A, Rosner I, Toubi E.
Clin Rev Allergy Immunol. 2008 Feb;34(1):74-9. **Review**.
PMID: 18240027 [PubMed - indexed for MEDLINE]
Related citations

Filters: Manage Filters

Results by year

Find related data
Database: Select
Find items

Search details
"Antibodies, Monoclonal"[Mesh] AND "Neoplasms"[Mesh] AND (Review[ptyp] AND English[lang])
Search See more...

Facts and knowledge (II)

Gene product annotations using Gene ontology

Breast cancer type 2 susceptibility protein

protein from *Homo sapiens* (human)

Term associations ⬇ Gene product information ➡ Peptide Sequence ➡ Sequence information ➡

Term Associations

Download all association information in: gene association format RDF-XML

Filter associations displayed

Select all | Clear all | Perform an action with this page's selected terms... | Go!

Accession, Term		Ontology	Qualifier	Evidence	Reference	Assigned by
GO:0007420 : brain development	3164 gene products view in tree	biological process		IEA With Ensembl:ENSMUSP00000038576	GO REF:0000019	Ensembl (via UniProtKB)
GO:0007569 : cell aging	838 gene products view in tree	biological process		IEA With Ensembl:ENSMUSP00000038576	GO REF:0000019	Ensembl (via UniProtKB)
GO:0033205 : cell cycle cytokinesis	1106 gene products view in tree	biological process		IDA	PMID:17286961	UniProtKB
GO:0051298 : centrosome duplication	319 gene products view in tree	biological process		IMP	PMID:17286961	UniProtKB
GO:0006978 : DNA damage response, signal transduction by p53 class mediator resulting in transcription of p21 class mediator	76 gene products view in tree	biological process		IEA With Ensembl:ENSMUSP00000038576	GO REF:0000019	Ensembl (via UniProtKB)

Facts and knowledge (III)

- The monthly cost of Herceptin is $4,500
- Herceptin is produced by Roche
- Herceptin has global sales of 5.25 billion Swiss francs in 2011
- Clinical trials (HERA, PHARE): One year on Herceptin is best
- Marie C. got heart failure after being treated with Herceptin

Sport | TV&Showbiz | Femail | Health | Science | Money | RightMinds

rectory | Health Boards | Diets | MyDish Recipe Finder

Cancer therapy saved my life but left me with heart failure

By JO WATERS

PUBLISHED: 00:23 GMT, 23 October 2012 | **UPDATED:** 00:23 GMT, 23 October 2012

Comments (15) | Share +1 0 Tweet 3 Like 17

When Marie Constantas was being treated for breast cancer, she got through it by focusing on the day she would be cured and could get on with life.

'I had chemotherapy to shrink my tumour, then a mastectomy and radiotherapy,' recalls Marie, 53, a former fabric technology specialist, who lives in Manchester.

'It killed my cancer, and then it was over.

'I was just incredibly grateful to be alive.'

But three weeks after she finished her radiotherapy treatment, Marie noticed something.

'I started feeling breathless when I walked upstairs.

'It was as if I had someone sitting on my shoulders weighing me down and every step required major effort,' she says.

'I'd never had any breathing or heart problems before — I was a slim size 10, ate well and kept fit.'

'I had chemotherapy to shrink my tumour, then a mastectomy and radiotherapy,' said Marie Constantas

Analyzing meaning: coded data

• ICD: C50.4	Neoplasm of upper-outer quadrant of breast
• ICD-O:M8500	Invasive ductal carcinoma
• TNM:T1N0M0	tumor 1.0 cm or less
• SCT:392021009	Lumpectomy of breast
• SCT:387003001	Trastuzumab
• LOINC:48676-1	HER2 in Tissue
• MeSH:D000911	Antibodies, Monoclonal
• MeSH:D009369	Neoplasms
• GO:0007569	cell aging
• GO:0006281	DNA repair
• GO:0005634	nucleus
• UniProt:P38398	Breast cancer type 1 susceptibility protein

Analyzing context

- Both coded content and sentences are context dependent
- Examples:

	[exists at coding]
Breast cancer (family history)	√
Breast cancer (hypothesis of GP, motivates referral)	?
Breast cancer (suspicious due to mammogram)	?
Breast cancer (confirmed fact after surgery)	√
Neoplasm (negated in previous history)	∅
chemotherapy (planned treatment)	∅
heart disorder (risk)	∅
monoclonal antibody (topic in scientific paper)	√
survival rate (estimated number according to cohort)	√

Observations

- In EHRs most clinical facts are encoded as free text narratives despite tendency towards more structured data
- Coding systems cover most concepts in health care and biomedical research
- Codes as well as textual statements are highly context-dependent
- Not everything exists (related to a patient) at the time a term, phrase, or code is used
- Use of codes (e.g. for DSS) requires further analysis of the underlying models of meaning

Types of models of meaning

- Thesauri / Terminologies
- Classifications
- Information models
- **Ontologies**

Medical Thesauri / Terminologies

- Groups together words / terms according to similarity in meaning
- Basic relations:
  - Synonymy
  - Broader / Narrower (ordering relations)
- Concept = Group of (quasi)synonyms
- Multiple hierarchies
- Mainly designed for retrieval
- Text definitions / explanations (scope notes) if required
- No formal semantics
- Medical terminologies
  - term standardization
  - controlled vocabulary
  - coding of clinical facts

Example: MeSH Thesaurus: trees

- Neoplasms [C04]
  - Neoplasms by Site [C04.588]
    - Abdominal Neoplasms [C04.588.033] +
    - Anal Gland Neoplasms [C04.588.083]
    - Bone Neoplasms [C04.588.149] +
    - ▶ Breast Neoplasms [C04.588.180]
      - Breast Neoplasms, Male [C04.588.180.260]
      - Carcinoma, Ductal, Breast [C04.588.180.390]
      - Hereditary Breast and Ovarian Cancer Syndrome [C04.588.180.483]
      - Inflammatory Breast Neoplasms [C04.588.180.576]
    - Digestive System Neoplasms [C04.588.274] +
    - Endocrine Gland Neoplasms [C04.588.322] +
    - Eye Neoplasms [C04.588.364] +
    - Head and Neck Neoplasms [C04.588.443] +
    - Hematologic Neoplasms [C04.588.448] +
    - Mammary Neoplasms, Animal [C04.588.531] +
    - Nervous System Neoplasms [C04.588.614] +
    - Pelvic Neoplasms [C04.588.699]
    - Skin Neoplasms [C04.588.805] +
    - Soft Tissue Neoplasms [C04.588.839] +
    - Splenic Neoplasms [C04.588.842]
    - Thoracic Neoplasms [C04.588.894] +
    - Urogenital Neoplasms [C04.588.945] +

- Skin and Connective Tissue Diseases [C17]
  - Skin Diseases [C17.800]
    - Breast Diseases [C17.800.090]
      - Breast Cyst [C17.800.090.249]
      - ▶ Breast Neoplasms [C17.800.090.500]
        - Breast Neoplasms, Male [C17.800.090.500.260]
        - Carcinoma, Ductal, Breast [C17.800.090.500.390]
        - Hereditary Breast and Ovarian Cancer Syndrome [C17.800.090.500.483]
        - Inflammatory Breast Neoplasms [C17.800.090.500.576]
      - Fibrocystic Breast Disease [C17.800.090.750]
      - Gynecomastia [C17.800.090.875]
      - Lactation Disorders [C17.800.090.937] +
      - Mastitis [C17.800.090.968] +

Example: Medical Subject Headings (MeSH)

Literature search: monoclonal antibodies and cancer therapy

PubMed.gov
US National Library of Medicine
National Institutes of Health

PubMed ("Antibodies, Monoclonal"[Mesh]) AND "Neoplasms"[Mesh] Search

MeSH terms

RSS Save search Advanced Help

Show additional filters

Clear all

Text availability
Abstract available
Free full text available
Full text available

Publication dates
5 years
10 years
Custom range...

Species
Humans
Other Animals

Article types clear
Clinical Trial
Meta-Analysis
Practice Guideline
Randomized Controlled Trial
✓ Review
Systematic Reviews
more ...

Display Settings: Summary, 20 per page, Sorted by Recently Added Send to:

Results: 2021 to 2040 of 5979

<< First < Prev Page 102 of 299 Next > Last >>

Filters activated: Review, English Clear all

2021. Trastuzumab as the lead monoclonal antibody in advanced breast cancer: choosing which patient and when.
Olver IN.
Future Oncol. 2008 Feb;4(1):125-31. **Review**.
PMID: 18241007 [PubMed - indexed for MEDLINE]
Related citations

2022. Predicting outcome to VEGF-targeted therapy in metastatic clear-cell renal cell carcinoma: data from recent studies.
Golshayan AR, Brick AJ, Choueiri TK.
Future Oncol. 2008 Feb;4(1):85-92. **Review**.
PMID: 18241003 [PubMed - indexed for MEDLINE]
Related citations

2023. Rituximab: beyond simple B cell depletion.
Kessel A, Rosner I, Toubi E.
Clin Rev Allergy Immunol. 2008 Feb;34(1):74-9. **Review**.
PMID: 18240027 [PubMed - indexed for MEDLINE]
Related citations

Filters: Manage Filters

Results by year

Find related data
Database: Select
Find items

Search details
```
"Antibodies, Monoclonal"[Mesh] AND
"Neoplasms"[Mesh] AND
(Review[ptyp] AND English[lang])
```
Search See more...

Example: MeSH Thesaurus: concept

MeSH Descriptor Data

Return to Entry Page

Standard View. Go to Concept View; Go to Expanded Concept View

MeSH Heading	Breast Neoplasms
Tree Number	C04.588.180
Tree Number	C17.800.090.500
Annotation	human only; BREAST NEOPLASMS, MALE is also available; for animal use MAMMARY NEOPLASMS, ANIMAL or MAMMARY NEOPLASMS, EXPERIMENTAL: Manual 24.5+, 24.6+; coordinate IM with histological type of neoplasm (IM)
Scope Note	Tumors or cancer of the human BREAST.
Entry Term	Breast Cancer
Entry Term	Breast Tumors
Entry Term	Cancer of Breast
Entry Term	Cancer of the Breast
Entry Term	Human Mammary Carcinoma
Entry Term	Mammary Carcinoma, Human
Entry Term	Mammary Neoplasm, Human
Entry Term	Mammary Neoplasms, Human
Entry Term	Neoplasms, Breast
Entry Term	Tumors, Breast
Allowable Qualifiers	BL BS CF CH CI CL CN CO DH DI DT EC EH EM EN EP ET GE HI IM ME MI MO NU PA PC PP PS PX RA RH RI RT SC SE SU TH UL UR US VE VI
Entry Version	BREAST NEOPL
Date of Entry	19990101
Unique ID	D001943

Example: MeSH Thesaurus: subheading

MeSH Qualifier Data

Return to Entry Page

Subheading	prevention & control
Record Type	Q
Entry Version	PREV
Abbreviation	PC
Scope Note	Used with disease headings for increasing human or animal resistance against disease (e.g., immunization), for control of transmission agents, for prevention and control of environmental hazards, or for prevention and control of social factors leading to disease. It includes preventive measures in individual cases.
Annotation	subhead only; includes "preventive therapy"; includes vaccination; for prevention on epidemiologic or personal level; see MeSH scope note in Introduction; indexing policy: Manual 19.8.58; DF: /prev or /PC CATALOG: may be subdivided geographically
Online Note	search policy: Online Manual; use: main heading/PC or PC (SH) or SUBS APPLY PC
History Note	66; used with Category C & F 1966-67; C, F & G 1968-74; C, F1, F3, G3 & I1 1975-89; C & F3 1990 forward
Entry Term	control
Entry Term	prevention
Entry Term	preventive measures
Entry Term	preventive therapy
Entry Term	prophylaxis
Date of Entry	19731227
Revision Date	20040713
Date Established	19960101
Unique ID	Q000517

Classifications

- System for organization of individuals into predefined containers (classes)
- Primary use: statistics oriented
- Extensional
- Enforced monohierarchical structure
- Disjoint classes
- Class labels are not terms
- No term definitions, rather classification instructions
- Residual classes ("not elsewhere classified", "not otherwise specified")

Ingenerf J, Giere W. Concept-oriented standardization and statistics-oriented classification: continuing the classification versus nomenclature controversy. Methods Inf Med. 1998 Nov;37(4-5):527-39.

Example: ICD-10

Malignant neoplasm of breast (C50-C50)

C50	**Malignant neoplasm of breast**
	Incl.: connective tissue of breast
	Excl.: skin of breast (C43.5, C44.5)
C50.0	**Nipple and areola**
C50.1	**Central portion of breast**
C50.2	**Upper-inner quadrant of breast**
C50.3	**Lower-inner quadrant of breast**
C50.4	**Upper-outer quadrant of breast**
C50.5	**Lower-outer quadrant of breast**
C50.6	**Axillary tail of breast**
C50.8	**Overlapping lesion of breast**
	[See note 5 at the beginning of this chapter]
C50.9	**Breast, unspecified**

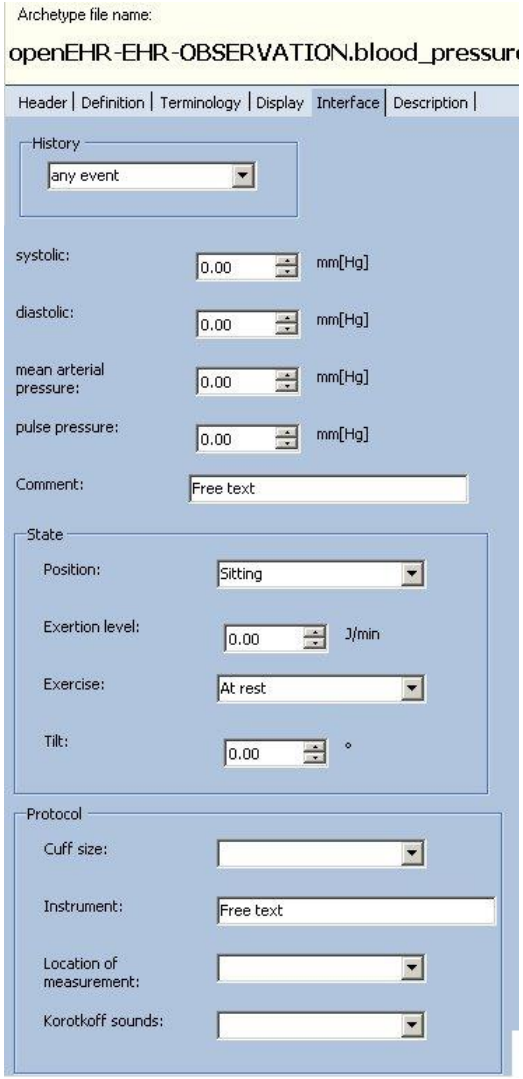
Information models

- Templates for structured acquisition of clinical data
- Standards:
  - HL7 (Clinical Models)
  - openEHR (Archetypes, Templates)
  - EN 13606 (Archetypes)
- Specify
  - Data elements
  - Value constraints
  - Vocabulary
  - "Bindings" to external terminology systems

Example: blood pressure archetype

```
archetype
   CEN-EHR-ENTRY.BloodPressure.v1
concept
   [at0000] -- Blood Pressure
definition
   ENTRY[at0000] matches {        -- Blood Pressure
      items cardinality matches {0..1; unordered} matches {
         ELEMENT[at0001] matches {   -- Systolic
            value matches {
               PHYSICAL_QUANTITY matches {
                  value matches {|10.0..500.0|}
                  property matches {"pressure"}
                  units matches {"mm[Hg]"}
               }
            }
         }
         ELEMENT[at0002] matches {   -- Dyastolic
            value matches {
               PHYSICAL_QUANTITY matches {
                  value matches {|10.0..500.0|}
                  property matches {"pressure"}
                  units matches {"mm[Hg]"}
               }
            }
         }
         ELEMENT[at0003] matches {   -- DateTime
            value matches {yyyy-mm-dd HH:MM:SS}
         }
      }
   }
```

```
ontology
   primary_language = <"en">
   languages_available = <"en","es">
   term_definitions = <
      ["en"] = <
         items = <
            ["at0000"] = <
               description = <"Blood Pressure measurement">
               text = <"Blood Pressure">
            >
            ["at0001"] = <
               description = <"Systolic Pressure">
               text = <"Systolic">
            >
            ["at0002"] = <
               description = <"Dyastolic Pressure">
               text = <"Dyastolic">
            >
            ["at0003"] = <
               description = <"Time of measurement">
               text = <"DateTime">
            >
         >
      >
```



Adolfo Muñoz, Roberto Somolinos, Mario Pascual, Juan A Fragua, Miguel A González, Jose Luis Monteagudo, Carlos H Salvador. Proof-of-concept Design and Development of an EN13606-based Electronic Health Care Record Service. J Am Med Inform Assoc 2007;14:118-129

Types of models of meaning

- Thesauri / Terminologies
- Classifications
- Information models
- **Ontologies**

(Formal) Ontologies

- *No universally accepted agreement of what ontologies are!*
- Computer science view
  - Ontologies are purpose oriented formal models of meaning (conceptualizations)
- Cognitive / linguistic views
  - Ontologies are concept systems or systems of semantic reference (no clear distinction from thesauri)
- Philosophy view (scientific realism)
  - Ontology is the study of what there is
  - Ontology is reality representation
  - Ontologies give precise mathematical formulations of the properties and relations of certain entities
  - Ontologies are built on domain-independent basic categories and ground axioms

(upper level ontology)

Quine O. On what there is. In: Gibson R. Quintessence - Basic Readings from the Philosophy of W. V. Quine. Cambridge: Belknap Press, Harvard University, 2004.
Schulz S, Stenzhorn H, Boeker M, Smith B: Strengths and limitations of formal ontologies in the biomedical domain. RECIIS - Electronic Journal in Communication, Information and Innovation in Health, 2009; 3 (1): 31-45:

Ontologies in biomedical science and health care

- Current mainstream (?) view
  - Ontologies are hierarchies of classes of domain entities
  - Ontologies are limited to describe what is universally true for all instances of a class
  - Ontologies subscribe to the model-theoretic semantics of description logics
  - Ontologies should be created using the Semantic Web standard OWL
- Contentious issues
  - Are ontologies kinds of knowledge representations?
    philosophical antinomy: *ὄντος* (being) vs. *ἐπιστήμη* (knowledge)
  - Class - instance (universal - individual) division given by nature or at the discretion of the ontology engineer?
  - Do ontologies require upper level models? Can and should they be domain-independent?

Example of Upper-level Ontology: BioTopLite

Top level classes

- Thing
  - particular
    - condition
    - disposition
    - 'immaterial object'
    - 'information object'
    - 'material object'
    - process
    - quality
    - role
    - time
    - 'value region'

Top level relations

- topObjectProperty
  - 'caused by'
  - causes
  - 'denoted by'
  - denotes
  - 'has duration'
  - 'has locus'
  - 'has origin'
  - 'has part'
  - 'has participant'
  - 'has point in time'
  - 'has quality'
  - 'locus of'
  - 'originates from'
  - 'part of'
  - 'participates in'
  - 'physically connected to'
  - 'preceded by'
  - precedes
  - 'quality located'
  - 'quality location of'
  - 'quality of'

Example axioms

- process
  - Class: process
  - process SubClassOf 'has participant' some Thing
  - DisjointClasses: disposition, 'immaterial object', 'information object', 'material object',
  - process SubClassOf 'part of' only process
  - process SubClassOf 'has part' only process
  - process SubClassOf particular
  - process SubClassOf ('has duration' some 'time interval') or ('has point in time' some 'point in time')

Elena Beißwanger, Stefan Schulz, Holger Stenzhorn and Udo Hahn
BioTop: An Upper Domain Ontology for the Life Sciences - A Description of its Current Structure, Contents, and Interfaces to OBO Ontologies
in Applied Ontology, Volume 3, Issue 4, Pages 205-212, IOS Press, Amsterdam, The Netherlands, December 2008

Large biomedical ontologies

- OBO Foundry
- SNOMED CT
- NCIT

Ontologies in biomedical science and health care

- Most current biomedical ontology projects commit to a simple variant of description logics (OWL-EL++). Theoretical background is set theory.
- Principal types of axioms in OWL ontologies
  - Taxonomies (is-a hierarchies):
    Every homo sapiens is a primate, every primate is a vertebrate
  - Aristotelian class definitions (Genus + Differentia)
    Viral hepatitis is equivalent to hepatitis that is caused by some virus population
  - Partonomies (part-of hierarchies)
    Every liver is part of some digestive system and every digestive system is part of some organism
  - Disjoint partitions
    Nothing is both a human and a chimpanzee

Taxonomies

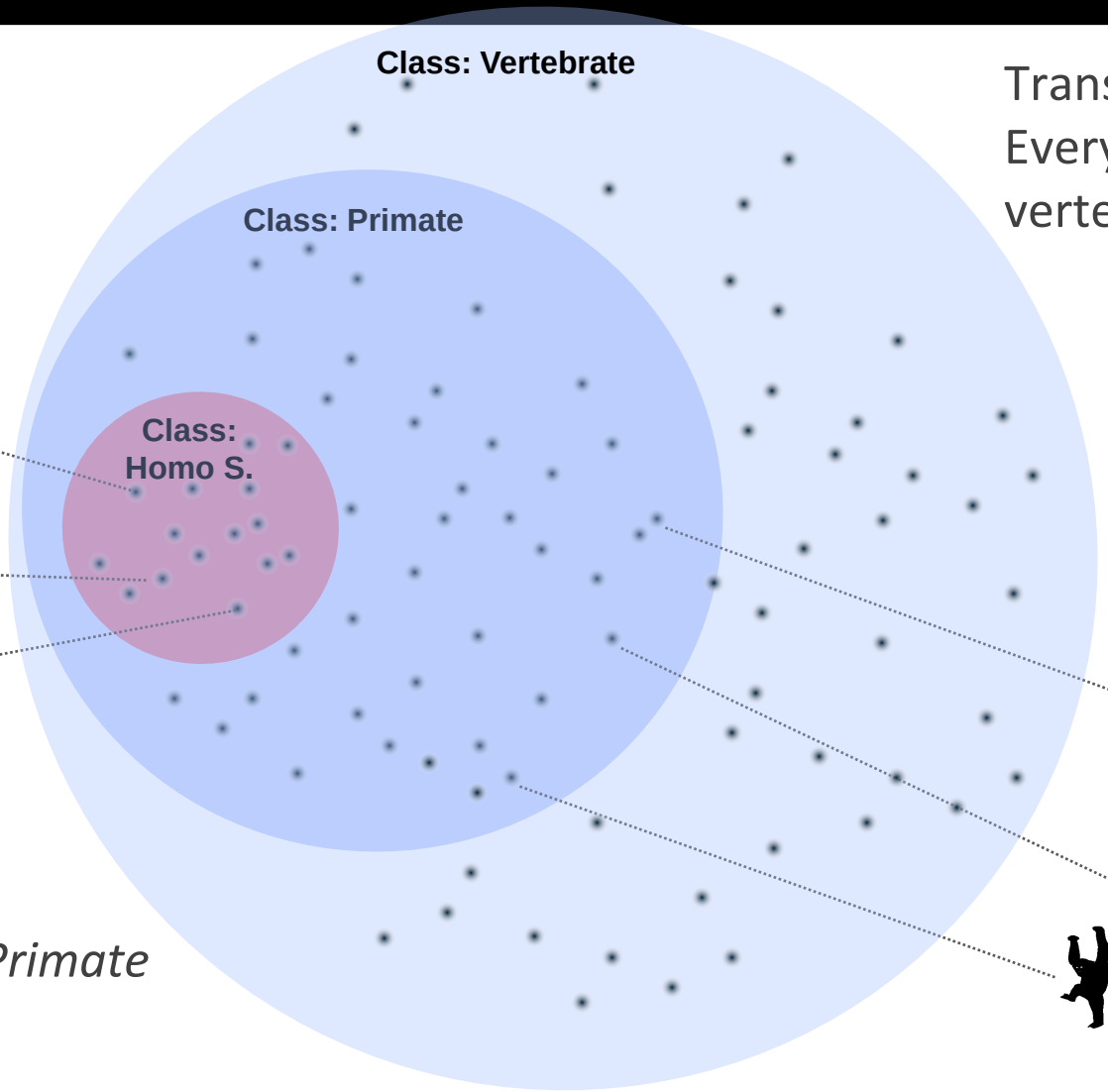
Every human is a primate, every primate is a vertebrate


Annette


Frank


Stefan

Human subClassOf *Primate*



Transitivity:
Every human is a vertebrate


Washoe


Koko

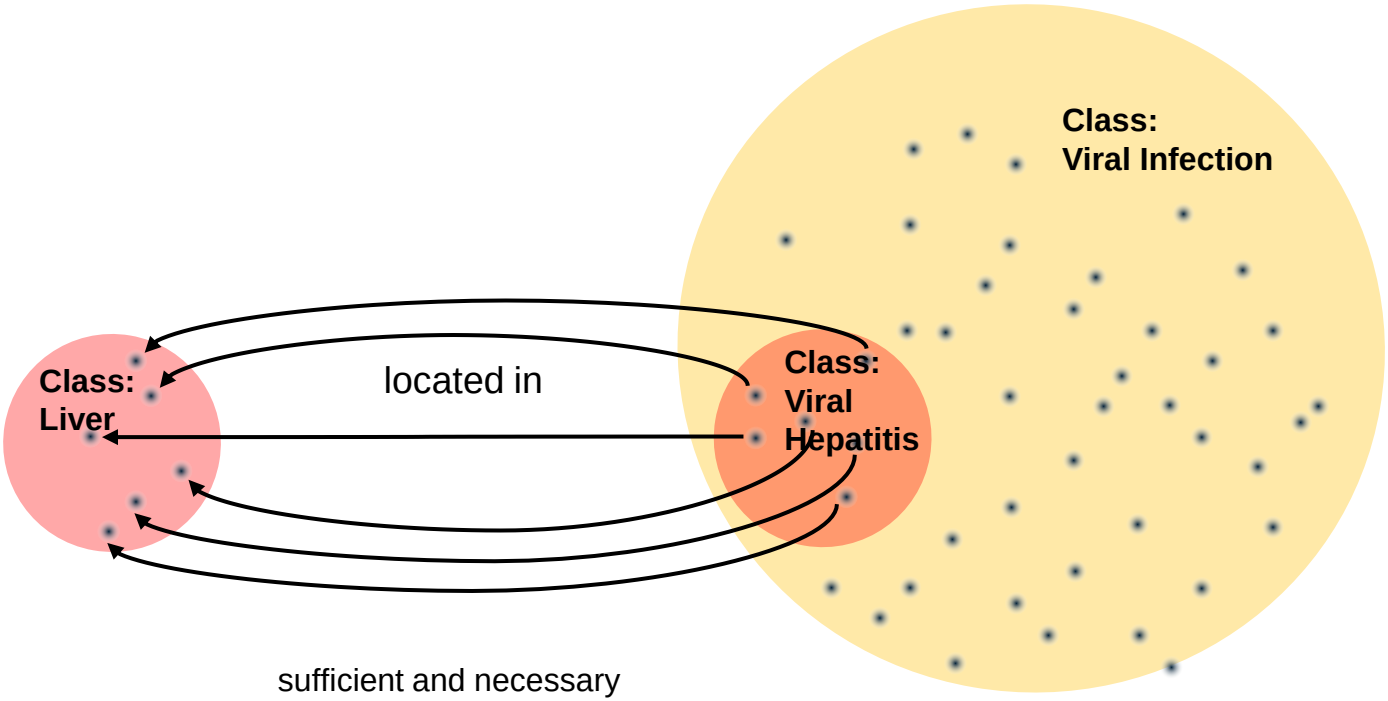
Bobo

Aristotelian definitions

Viral Hepatitis is equivalent to Viral infection that is located in some Liver

ViralHepatitis equivalentTo *ViralInfection* and **locatedIn** some *Liver*

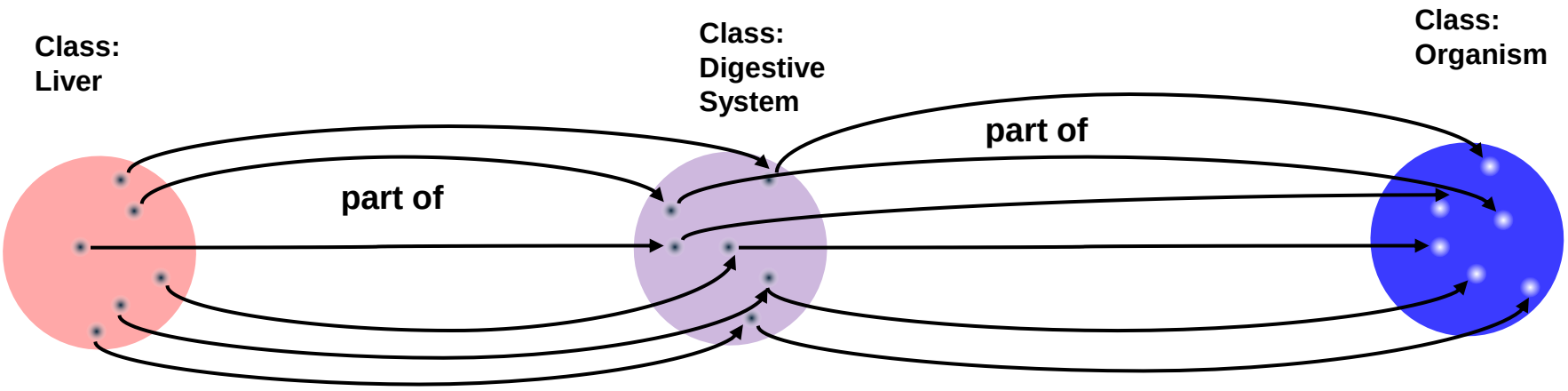
"species" *"genus"* *"differentia specifica"*



sufficient and necessary conditions

Partonomies

- ***Every liver is part of some digestive system and every digestive system is part of some organism***
- ***Transitivity: Every liver is part of some organism***



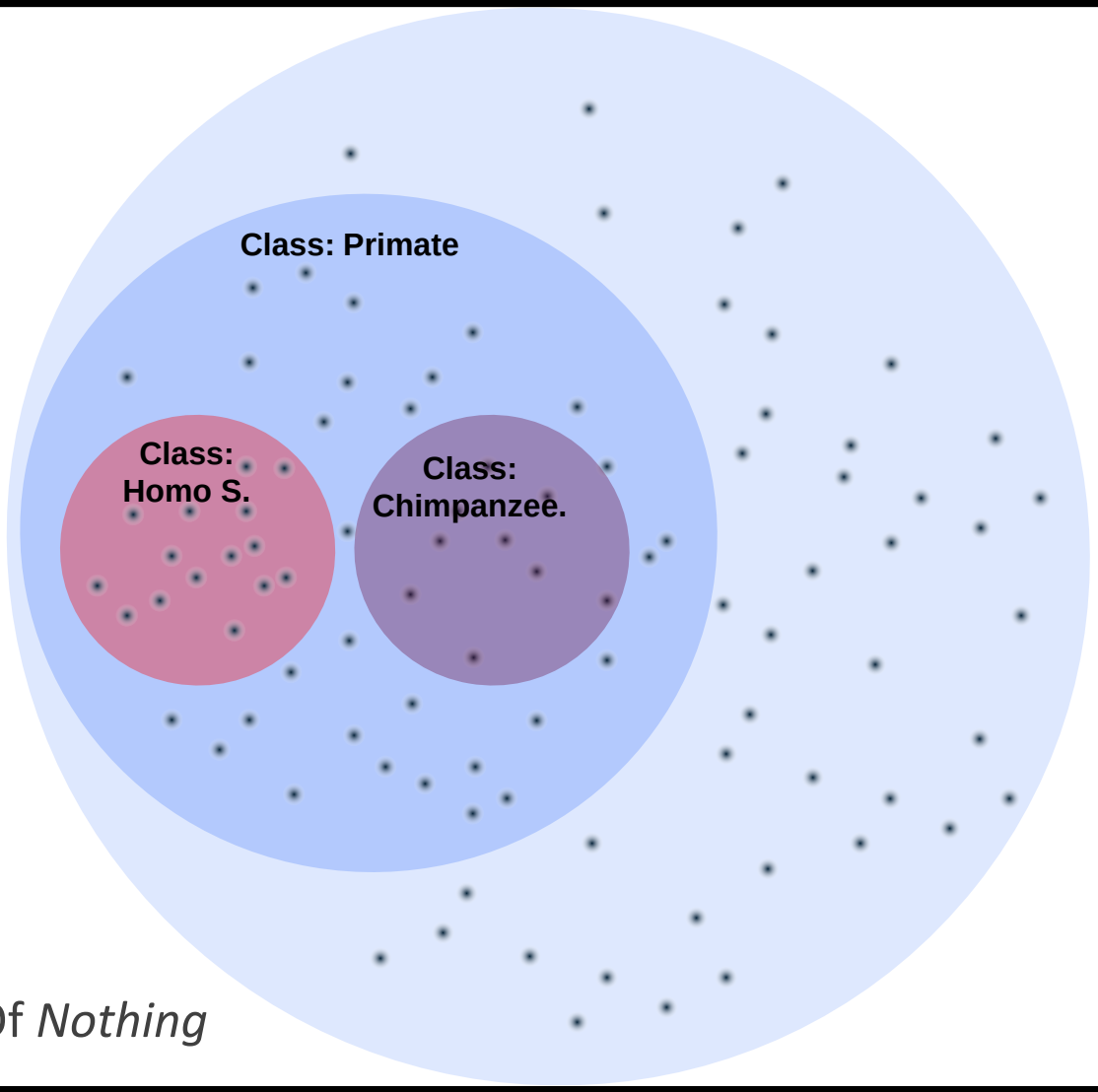
Liver subClassOf **partOf** some *DigestiveSystem*

DigestiveSystem subClassOf **partOf** some *Organism*

partOf o **partOf** subPropertyOf **partOf**

Disjoint partitions

No human is a chimpanzee



Human and *Chimp* subClassOf *Nothing*

The OBO Foundry

- Collaborative, user-driven, bottom up initiative
- Driven by the success of the Gene Ontology
- Guided by ontology development principles
- Rooted in upper ontologies (BFO + RO)
- Goal of creating a suite of orthogonal interoperable reference ontologies in the biomedical domain
- Division of labor amongst domain experts
- Originally using semi-formal OBO syntax, now increasingly OWL-DL

Smith, B.; Ashburner, M.; Rosse, C.; Bard, J.; Bug, W.; Ceusters, W.; Goldberg, L. J.; Eilbeck, K. et al. (2007). "The OBO Foundry: Coordinated evolution of ontologies to support biomedical data integration".Nature Biotechnology 25 (11): 1251–1255. doi:10.1038/nbt1346

OBO Foundry: Orthogonality by Upper-level, and Granularity divisions

RELATION TO TIME / GRANULARITY	CONTINUANT				OCCURRENT
	INDEPENDENT		DEPENDENT		
ORGAN AND ORGANISM	Organism (NCBI Taxonomy)	Anatomical Entity (FMA, CARO)	Organ Function (FMP, CPRO)	Phenotypic Quality (PaTO)	**Biological Process** (GO)
CELL AND CELLULAR COMPONENT	Cell (CL)	**Cellular Component** (FMA, GO)	Cellular Function (GO)		
MOLECULE	Molecule (ChEBI, SO, RnaO, PrO)		**Molecular Function** (GO)		Molecular Process (GO)

Linking OBO Foundry ontologies

- OBO Foundry principles
  - Single *is_a* parenthood, further *is_a* relations *inferred*
  - Genus-species definitions
- Reality: most Foundry ontologies have no equivalence axioms
- Example:
  - *Calcitonin secreting cell* (Cell Ontology) can be defined as a *Secretory cell* which secretes *Calcitonin* (ChEBI)
  - *Heart development* (Gene Ontology) can be defined as a *Developmental process* which has *Heart* (FMA) as participant
- Conclusion: Sufficient representation of definitional knowledge requires cross-linking OBO modules.

SNOMED CT (Systematized Nomenclature of Medicine - Clinical Terms)

INTERNATIONAL HEALTH TERMINOLOGY
STANDARDS DEVELOPMENT ORGANISATION

About IHTSDO | SNOMED CT | Join Us | News | Projects | Publications | Members | Contact Us | Home

Search Print FAQ

- About SNOMED CT
- Value Proposition for SNOMED CT
- SNOMED CT Lookup Service
- Release of SNOMED CT
- Licensing
- Namespaces
- History of SNOMED CT
- Request Submission
- Member Exchange
- Quality Assurance
- Tooling

SNOMED CT

SNOMED CT (**S**ystematized **No**menclature of **Med**icine-**C**linical **T**erms) is considered to be the most comprehensive, multilingual clinical healthcare terminology in the world. Each year, avoidable deaths and injuries occur because of poor communication between healthcare practitioners, or because busy practitioners forget or neglect to follow their own criteria for best practices. The delivery of a standard clinical terminology for use across the world's health information systems can therefore make a significant contribution towards improving the quality and safety of healthcare.

SNOMED CT aims to contribute to the improvement of patient care through underpinning the development of systems to accurately record health care encounters and to deliver decision support to health care providers. Ultimately, patients will benefit from the use of SNOMED CT to more clearly describe and accurately record their care, in building and facilitating better communication and interoperability in electronic health record exchange, and in creating systems that support health care decision making.

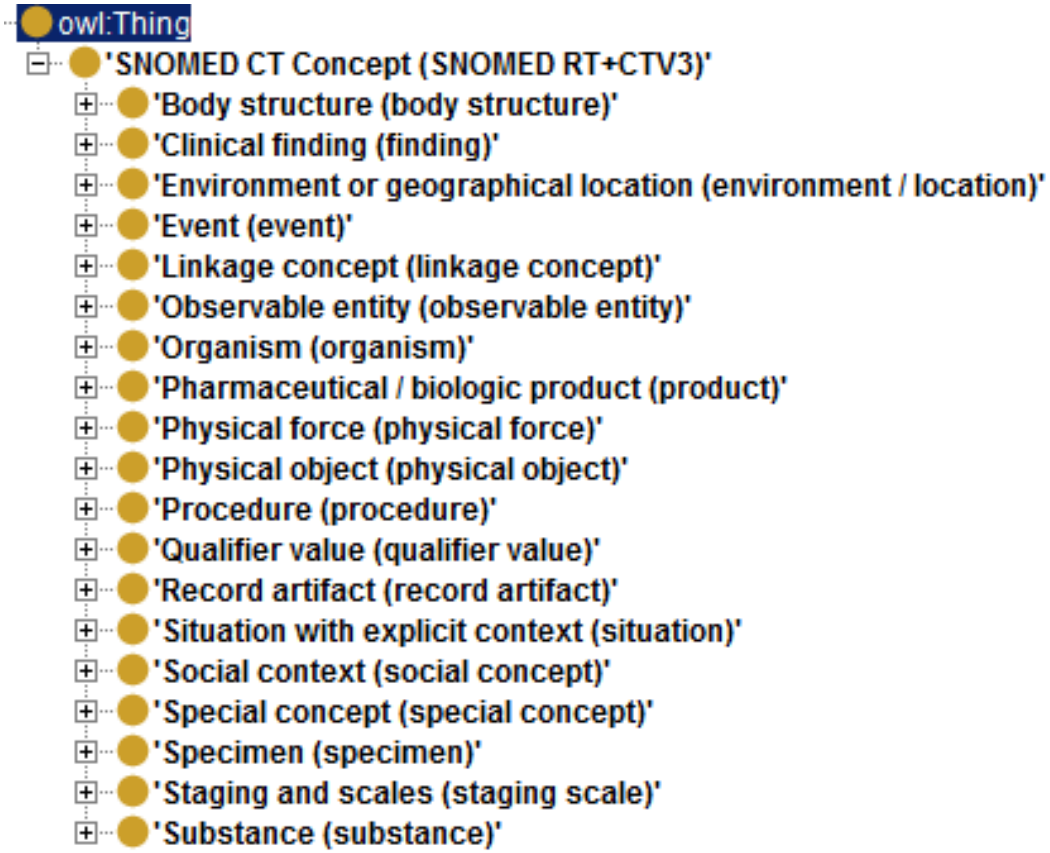
SNOMED CT intellectual property rights were transferred to the SNOMED SDO® in the formal creation of the IHTSDO.

SNOMED CT was originally created by the College of American Pathologists by combining SNOMED RT and a computer based nomenclature and classification known as Clinical Terms Version 3, formerly known as Read Codes Version 3, which was created on behalf of the UK Department of Health and is Crown copyright.

http://www.ihtsdo.org/

SNOMED CT - clinical terminology with ontological foundations

- Terminology for clinical data covering diseases, findings, procedures, organisms, substances etc.
- 311, 000 concepts, connected by 1,360,000 relational expressions
- advertised as a "terminological standard"



SNOMED CT: Terminology + Ontology

Concepts (represent-ational units)

Parent(s):
(Select a parent to make it the "Current Concept".)
Malignant neoplasm of thorax (disorder)
Neoplasm of breast (disorder)

Current Concept:
Malignant tumor of breast (disorder)

Child(ren):
(N=16) (Select a child to make it the "Current Concept".)
Carcinoma of breast (disorder)
Familial cancer of breast (disorder)
Hormone receptor positive malignant neoplasm of breast (disorder)
Local recurrence of malignant tumor of breast (disorder)
Malignant lymphoma of breast (disorder)
Malignant melanoma of breast (disorder)
Malignant neoplasm of axillary tail of breast (disorder)
Malignant neoplasm of breast lower inner quadrant (disorder)
Malignant neoplasm of breast lower outer quadrant (disorder)
Malignant neoplasm of breast upper inner quadrant (disorder)
Malignant neoplasm of breast upper outer quadrant (disorder)
Malignant neoplasm of female breast (disorder)
Malignant neoplasm of male breast (disorder)
Primary malignant neoplasm of breast (disorder)
Sarcoma of breast (disorder)

Axioms

Current Concept:
Fully Specified Name: Malignant tumor of breast (disorder)
ConceptId: 254837009

Defining Relationships:
Is a Malignant neoplasm of thorax (disorder)
Is a Neoplasm of breast (disorder)
Group 1
Associated morphology Malignant neoplasm of primary, secondary, or uncertain origin (morphologic abnormality)
Finding site Breast structure (body structure)
This concept is fully defined.

Qualifiers:
View Qualifying Characteristics and Facts

Descriptions (Synonyms):
Fully Specified Name: Malignant tumor of breast (disorder)
Preferred: Malignant tumor of breast [379661016]
Synonym: Breast cancer [379662011]
Synonym: CA - Breast cancer [379663018]
Preferred: Malignant tumour of breast [379664012]

Related Concepts:
- All "Is a" antecedents -
- All descendents/subtypes -
- Related concepts demo -

Terms

The NCIthesaurus

National Cancer Institute | U.S. National Institutes of Health | www.cancer.gov

EVS Enterprise Vocabulary Services | NCI Term Browser

NCIthesaurus

Search | ? | Exact Match | Begins With | Contains | Name/Code | Property | Relationship

Home | View Hierarchy | Subsets | Help

Quick Links

Welcome

Version: June 27, 2011 (11.06d)

NCI Thesaurus (NCIt) provides reference terminology for many NCI and other systems. It covers vocabulary for clinical care, translational and basic research, and public information and administrative activities.

NCIt features:

- Stable, unique codes for biomedical concepts;
- Preferred terms, synonyms, definitions, research codes, external source codes, and other information;
- Links to NCI Metathesaurus and other information sources;
- Over 200,000 cross-links between concepts, providing formal logic-based definition of many concepts;
- Extensive content integrated from NCI and other partners, much available as separate NCIt subsets
- Updated frequently by a team of subject matter experts.

NCIt is a widely recognized standard for biomedical coding and reference, used by a broad variety of public and private partners both nationally and internationally including the Clinical Data Interchange Standards Consortium Terminology (CDISC), the U.S. Food and Drug Administration (FDA), the Federal Medication Terminologies (FMT), and the National Council for Prescription Drug Programs (NCPDP).

Source Home Page: http://ncit.nci.nih.gov

Download: http://evs.nci.nih.gov/ftp1/NCI_Thesaurus

caBIG VKC Link: https://cabig-kc.nci.nih.gov/Vocab/KC/index.php/NCI_Thesaurus

EVS — NCI Enterprise Vocabulary Services: Terminology resources and services for NCI and the biomedical community.

NCIm — NCI Metathesaurus: Comprehensive database of 3,600,000 terms from 76 terminologies.

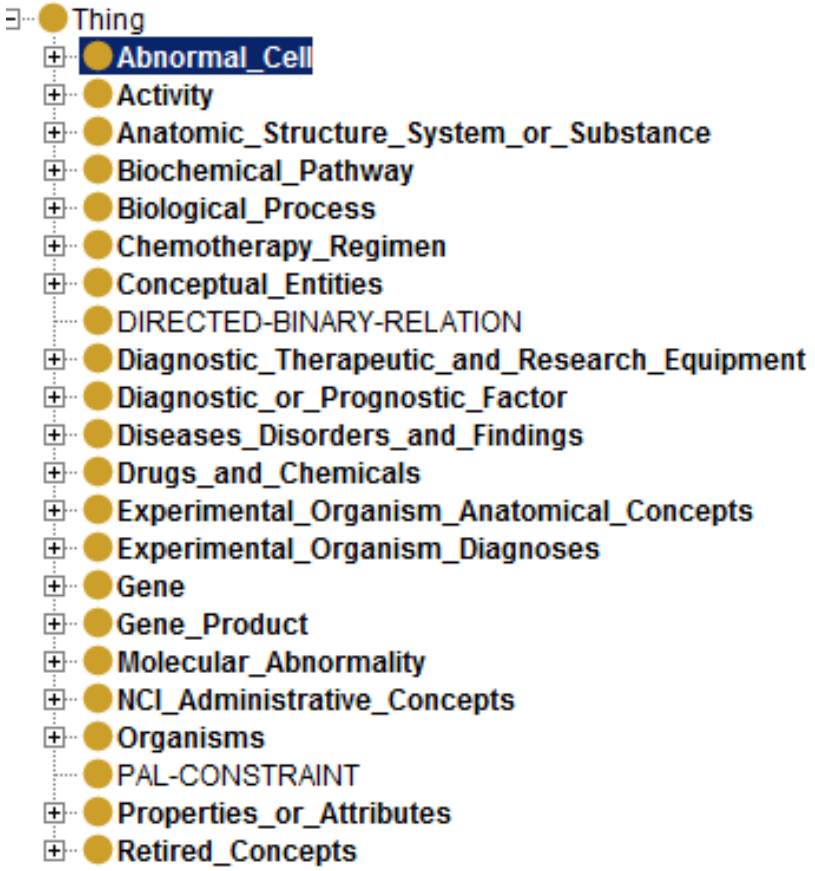
NCI Term Browser — NCI Term Browser: NCI and other terminologies in an integrated environment.

cancer.gov — NCI Terminology Resources: More information on NCI dictionaries and resources.

http://ncit.nci.nih.gov/ncitbrowser/

The NCI Thesaurus (NCIT) – a (N)ontology for cancer research

- Reference terminology of the National Cancer Institute (clinical care, translational and basic research, public and administrative information)
- 90k classes, 11k eq axioms, 110k subclass axioms, 23k classes interpreted as individuals ("punning") for enabling non-quantified triplets
- Upper-level partition (disjoint axioms at 1st hierarchical level)
- Distributed in OWL (SH(D)) as stated and inferred version
- But: "*NCI Thesaurus has some ontology-like features but NCI Thesaurus is not an ontology and is not designed or intended to one*."



General observations on SNOMED CT and NCIT

- Thesaurus / Ontology hybrids
  - SNOMED CT: ontology-aware redesign projects underway, increasing use of OWL for prototyping
  - NCIT: "no ontology": uses OWL syntax, but does not subscribe to OWL-DL semantics
- Both OWL versions are logically consistent
- DL reasoning used in design process, but not really in applications
- Numerous entailments are unreliable, due to plain design errors, but also due to "workarounds" to express negation or probabilistic knowledge

Typical category confusions

- Instances instead of subclasses
  - „*Insulin* Type *Peptide*“
- Superclasses instead of roles
  - *Fish* subClassOf *Food*
- Epistemic intrusion
  - *Infection of unknown origin* subClassOf *Infection*
- Hidden ambiguity
  - *Tumor* subClassOf *Pathological Process*
  - *Tumor* subClassOf *Pathological Body Part*
- Confusion function / process
  - *ATP transport* subClassOf *Biological Function*
  - *ATP transport* subClassOf *Biological Process*
- Confusion Process / Plan
  - *Planned Tonsillectomy* subClassOf *Tonsillectomy*
- Confusion material object / information object
  - *Thorax XRay* subClassOf **hasPart** some *Heart*

Other problems (source SNOMED)

- Anatomy-related entailments:

  AmputationOfTheFoot equivalentTo *rg* some
  (**method** some *Amputation* *and* **procedureSiteDirect** some *FootStructure)*

  AmputationOfToe *equivalentTo rg* some
  (**method** *some Amputation* and **procedureSiteDirect** some *ToeStructure)*

  ToeStructure subClassOf *FootStructure*

  AmputationOfToe subclassOf *AmputationOfTheFoot*

- Negations

  ThumbAbsent subClassOf *VenousFinding* and
  hasFindingSite some *ThumbStructure*

- Confusion Epistemology / Ontology

  PresumedViralAgent subClassOf *Virus*

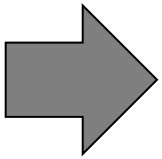
Schulz S, Suntisrivaraporn B, Baader F, Boeker M: SNOMED reaching its adolescence: Ontologists' and logicians' health check.
Int J Med Inform, 2009; 78 Suppl. 1: 86-94:http://dx.doi.org/10.1016/j.ijmedinf.2008.06.004

The translation of OAV triples into DL statements is ambiguous

Expressions like C1 Rel C2 are syntactically incorrect in OWL DL!

Translation of triples

C1 Rel C2



C1 subClassOf **rel** some *C2*

or

C1 subClassOf **rel** only *C2*

or

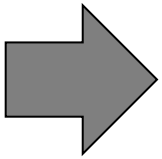
C2 subclassOf inv(**rel**) some *C2*

or…

Translation of groups of triples

C1 Rel C2

C1 Rel C3



C1 subClassOf (**rel** some *C2*) and (**rel** some *C3*)

or

C1 equivalentTo (**rel** some *C2*) and (**rel** some *C3*)

or

C1 equivalentTo (**rel** some ((*C2* or *C3*)))

or…

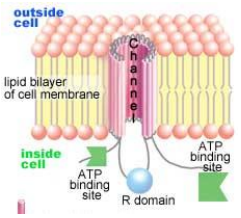
Typical problems when converting a thesaurus into an ontology

Mismatch between the intended meaning of labels and DL semantics:

- *Ureter_Small_Cell_Carcinoma* subclassOf
  Disease_May_Have_Finding some *Pain*

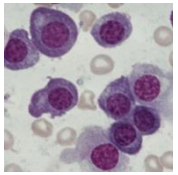
Incorrect distributive statements

- *Calcium-Activated_Chloride_Channel-2* subClassOf
  Gene_Product_Expressed_In_Tissue some *Lung* and
  Gene_Product_Expressed_In_Tissue some *Mammary_Gland* and
  Gene_Product_Expressed_In_Tissue some *Trachea*



Existential quantification over parts instead of wholes

- Antibody_Producing_Cell *subclassOf*
  Part_Of *some* Lymphoid_Tissue



Schulz S, Schober S, Tudose I, Stenzhorn H: The Pitfalls of Thesaurus Ontologization – the Case of the NCI Thesaurus. AMIA Annu Symp Proc, 2010: 727-731: http://proceedings.amia.org/127gtf/1

Ontologies are not exactly made for represent contingent knowledge

- *“Smoking Causes Lung Cancer” / "Aspirin treats Headache"*
- Interpretation problem:
  - Not every smoking (event, habit) causes some lung cancer
  - Not every lung cancer is caused by smoking
  - Not every aspirin tablet treats some headache
  - Not every headache is treated by some aspirin tablet
- Description logics do not allow probabilistic, default, or normative assertions
- Axioms can only state what is true for all members of a class
- Introducing dispositions into ontology possible but not very intuitive
  - every participant of a smoking event has some inherent disposition which is only realized by getting lung cancer
  - every aspirin tablet has the disposition of treating headache when ingested

Large biomedical ontologies are slowly maturing

- Legacy: originally not conceived as ontologies but as systems for classification and semantic annotation
- "Ontologization" of thesaurus-like structures problematic
- Few use cases for logic-based reasoning
- General tendency towards OWL
- Persisting problems
  - Understanding foundations of logic
  - Tendency to create idiosyncratic, non-principled models
  - Acceptance of upper-level ontology still insufficient
  - Context dependence of ontology statements not explicit
  - Limited interoperability between different ontologies

Ontology vs. Knowledge representation

- Ontologies describe *classes of real things (individuals)*, which exist in the world
- A natural divide between ontology and KR in general is the philosophical distinction between ontology and epistemology:
  1. Ontologies state what (is known to be) universally true for all members of a class, independent of observational context
  2. Knowledge representation artefacts additionally make contingent, probabilistic, fuzzy, or default statements on individuals or classes of individuals
- The use of ontology language should be restricted to 1.

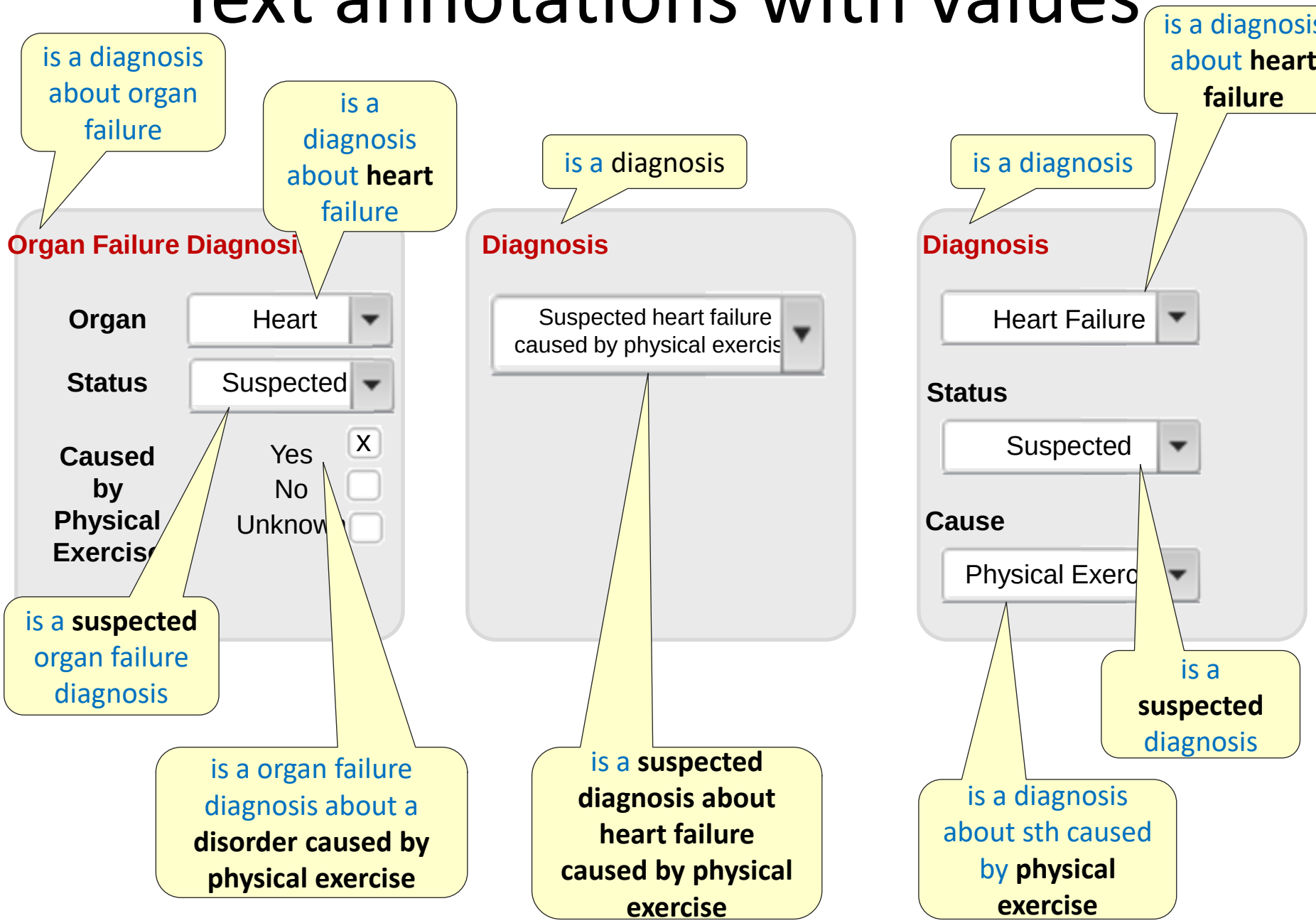
"There are very few interesting items of knowledge that are truly ontological in this strict sense"

Alan Rector

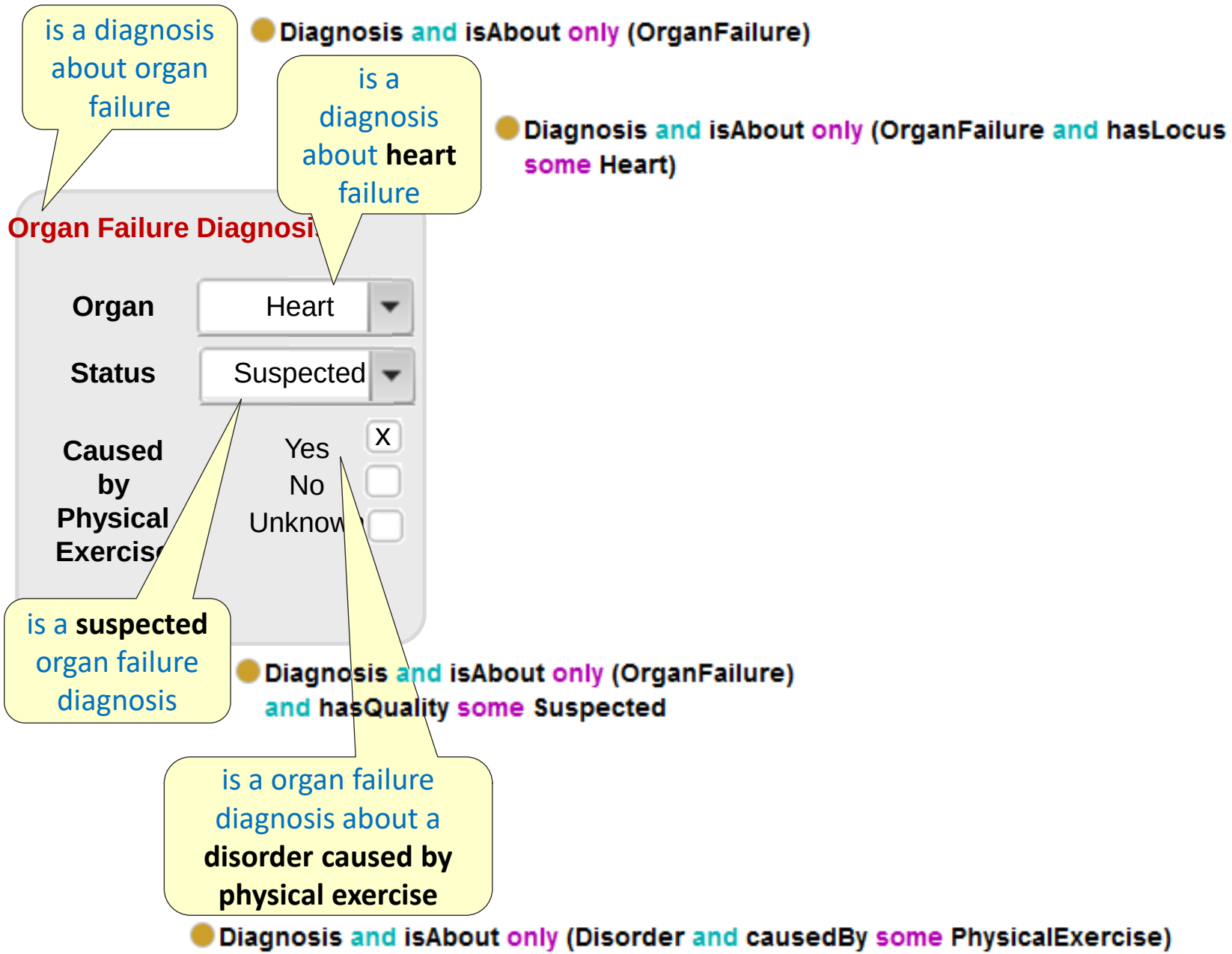
Rector A. Barriers, approaches and research priorities for integrating biomedical ontologies. Semantic Health Net deliverable 6.1. http://www.semantichealth.org/DELIVERABLES/SemanticHEALTH_D6_1.pdf

Use case: semantic annotation of information models + values

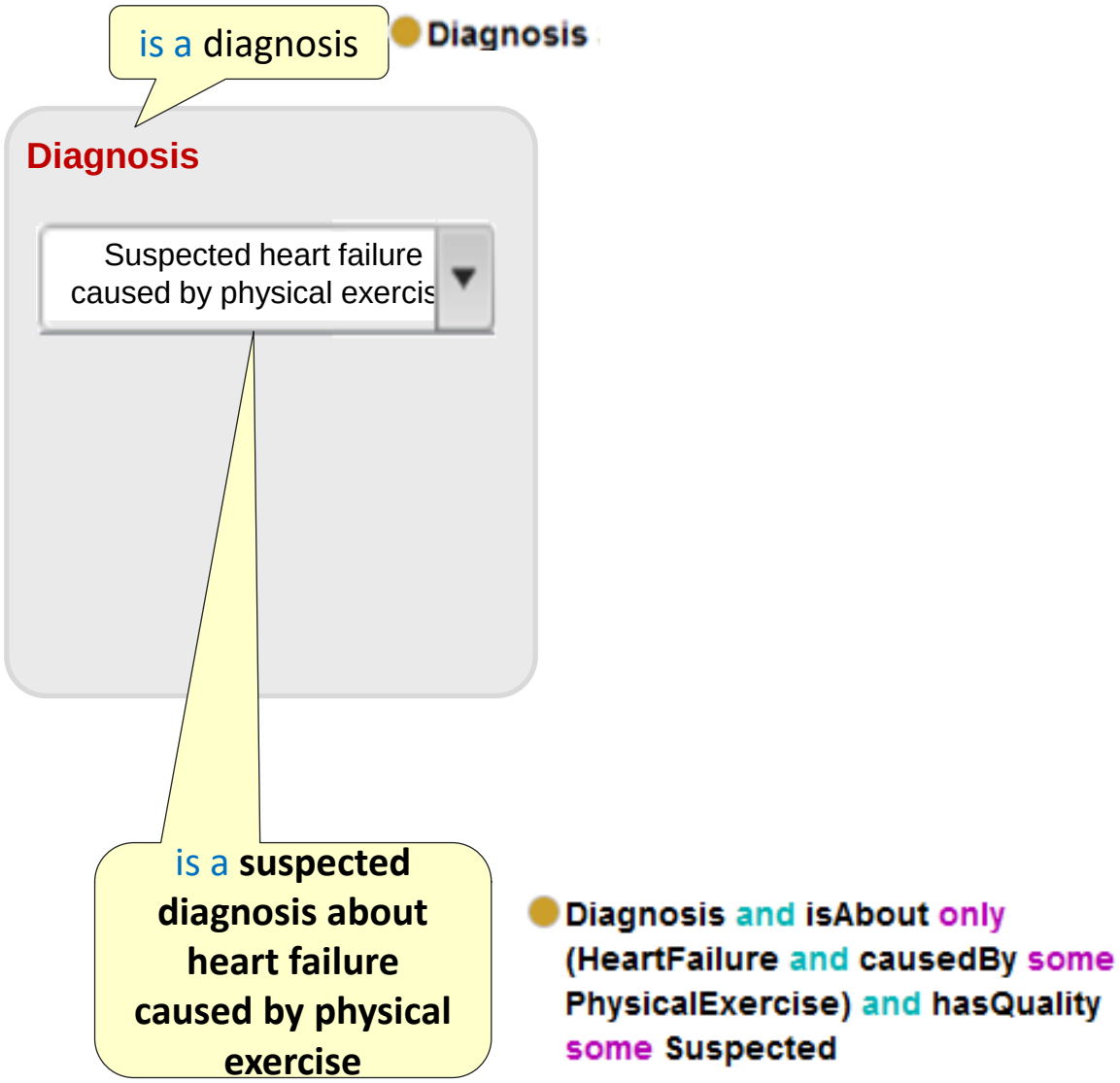
Text annotations with values



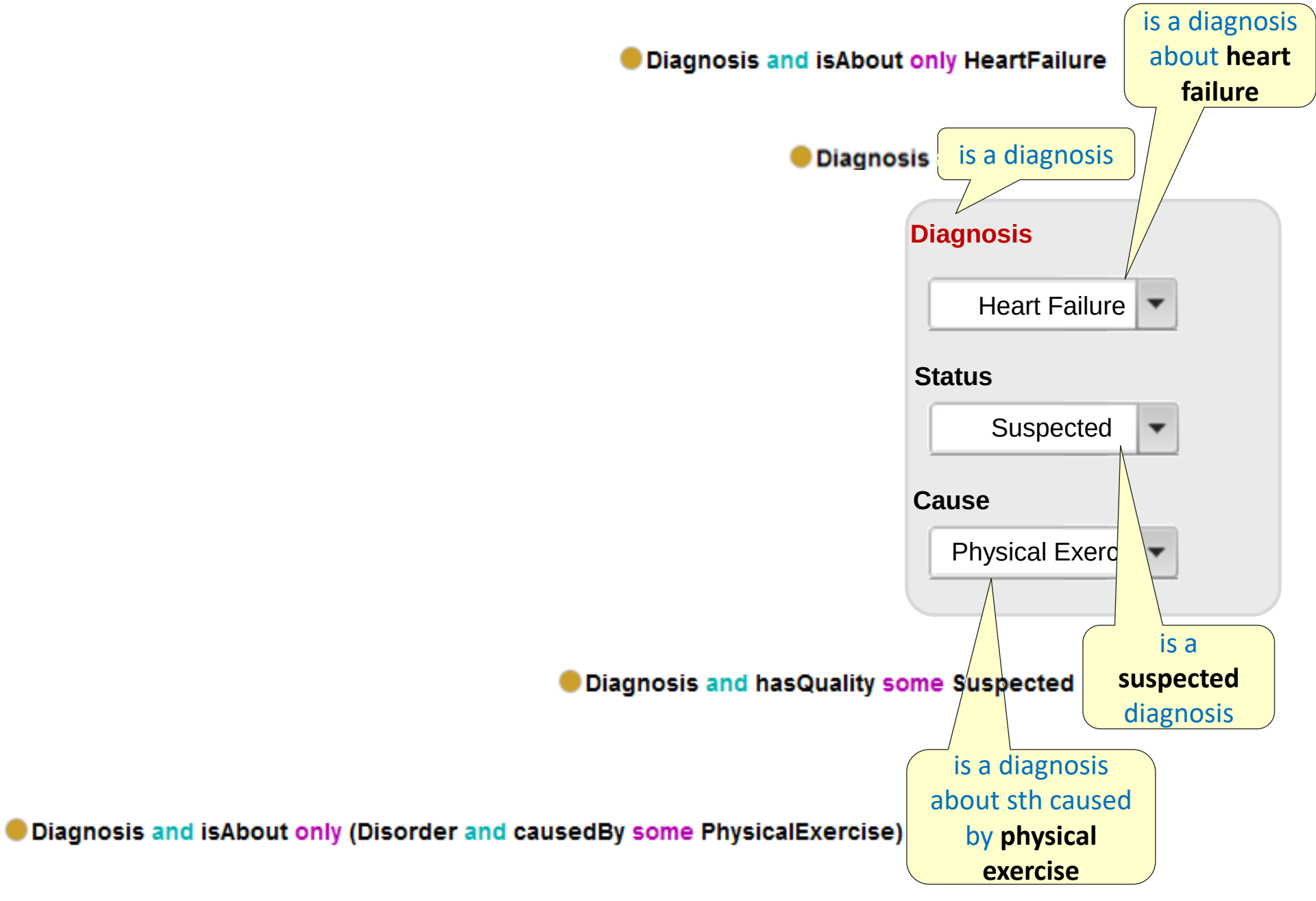
DL annotations



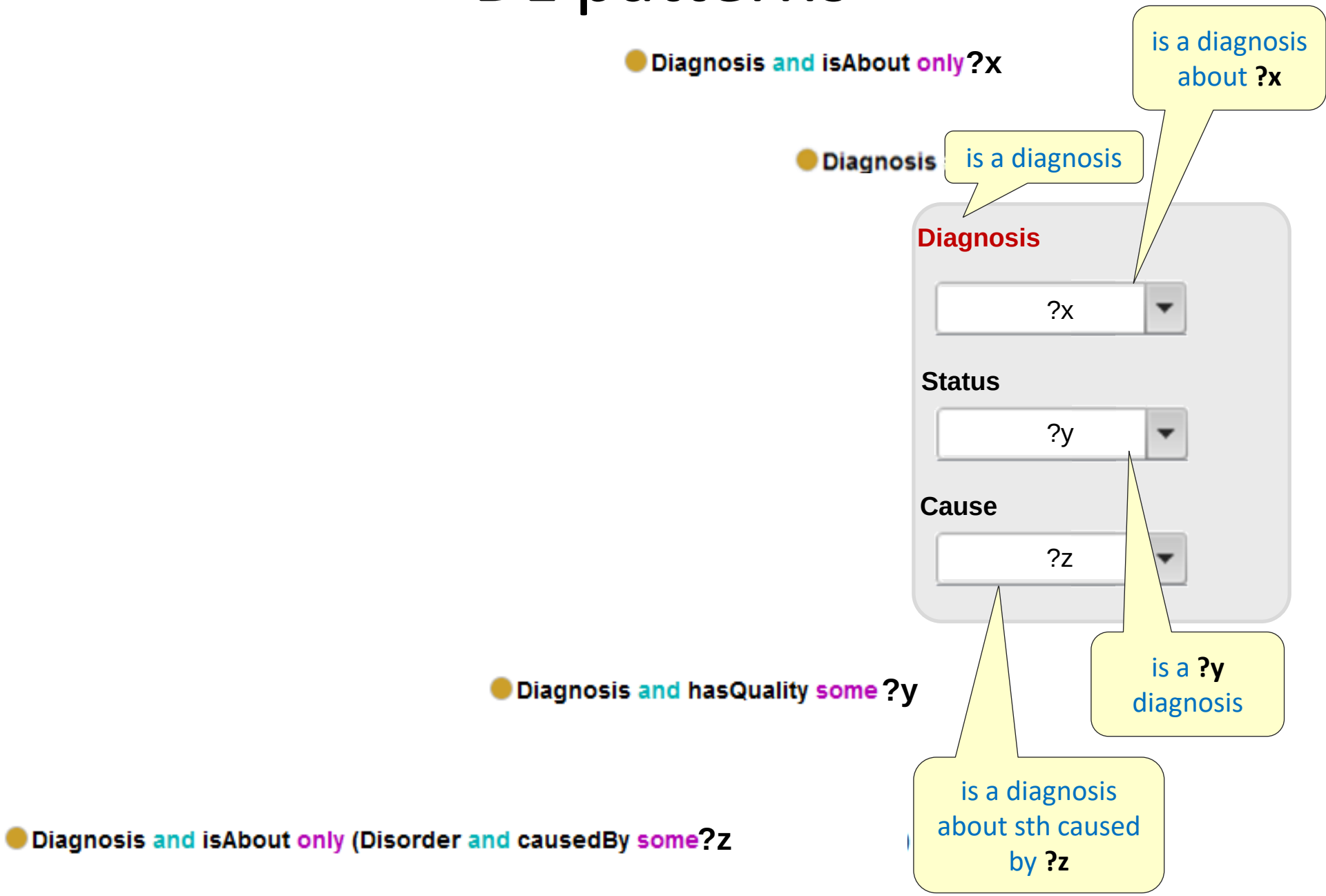
DL annotations



DL annotations



DL patterns


Diagnosis and isAbout only ?x
Diagnosis
Diagnosis and hasQuality some ?y
Diagnosis and isAbout only (Disorder and causedBy some ?z
is a diagnosis about ?x
is a diagnosis
Diagnosis
?x
Status
?y
Cause
?z
is a ?y diagnosis
is a diagnosis about sth caused by ?z

One diagnosis instance of each model

- Diagnosis_2 **Type** Diagnosis **and** (hasQuality **some** Suspected) **and** (isAbout **only** (HeartFailure **and** (causedBy **some** PhysicalExercise)))
- Diagnosis_2 **Type** Diagnosis

Organ Failure Diagnosis

Organ Heart

Status Suspected

Caused by Physical Exercise

Yes [X]
No [ ]
Unknown [ ]

Diagnosis

Suspected heart failure caused by physical exercis

Diagnosis

Heart Failure

Status

Suspected

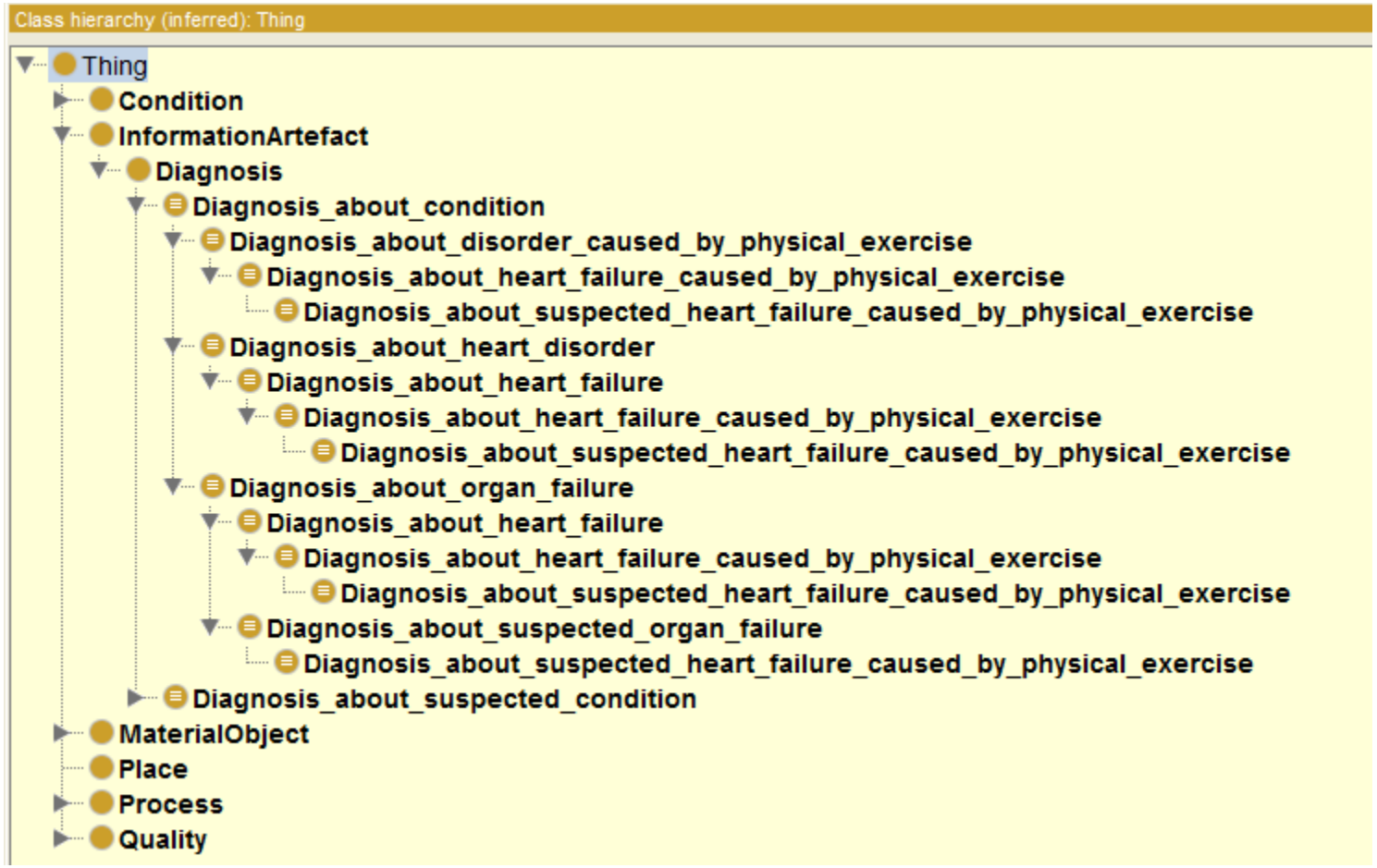
Cause

Physical Exercis

- Diagnosis_1 **Type** Diagnosis **and** (isAbout **only** OrganFailure)
- Diagnosis_1 **Type** Diagnosis **and** (hasQuality **some** Suspected)
- Diagnosis_1 **Type** Diagnosis **and** (isAbout **only** (Disorder **and** (causedBy **some** PhysicalExercise)))
- Diagnosis_1 **Type** Diagnosis **and** (isAbout **only** (OrganFailure **and** (hasLocus **some** Heart)))
- Diagnosis_1 **Type** Diagnosis

- Diagnosis_3 **Type** Diagnosis **and** (isAbout **only** HeartFailure)
- Diagnosis_3 **Type** Diagnosis **and** (isAbout **only** (Disorder **and** (causedBy **some** PhysicalExercise)))
- Diagnosis_3 **Type** Diagnosis **and** (hasQuality **some** Suspected)
- Diagnosis_3 **Type** Diagnosis

DL classification



DL querying (I)

Query (class expression)

Diagnosis **and** isAbout **only** (HeartFailure **and** causedBy **some** PhysicalExercise) **and** hasQuality **some** Suspected

Execute | Add to ontology

Query results

Equivalent classes (1)

- **Diagnosis_about_suspected_heart_failure_caused_by_physical_exercise**

Ancestor classes (11)

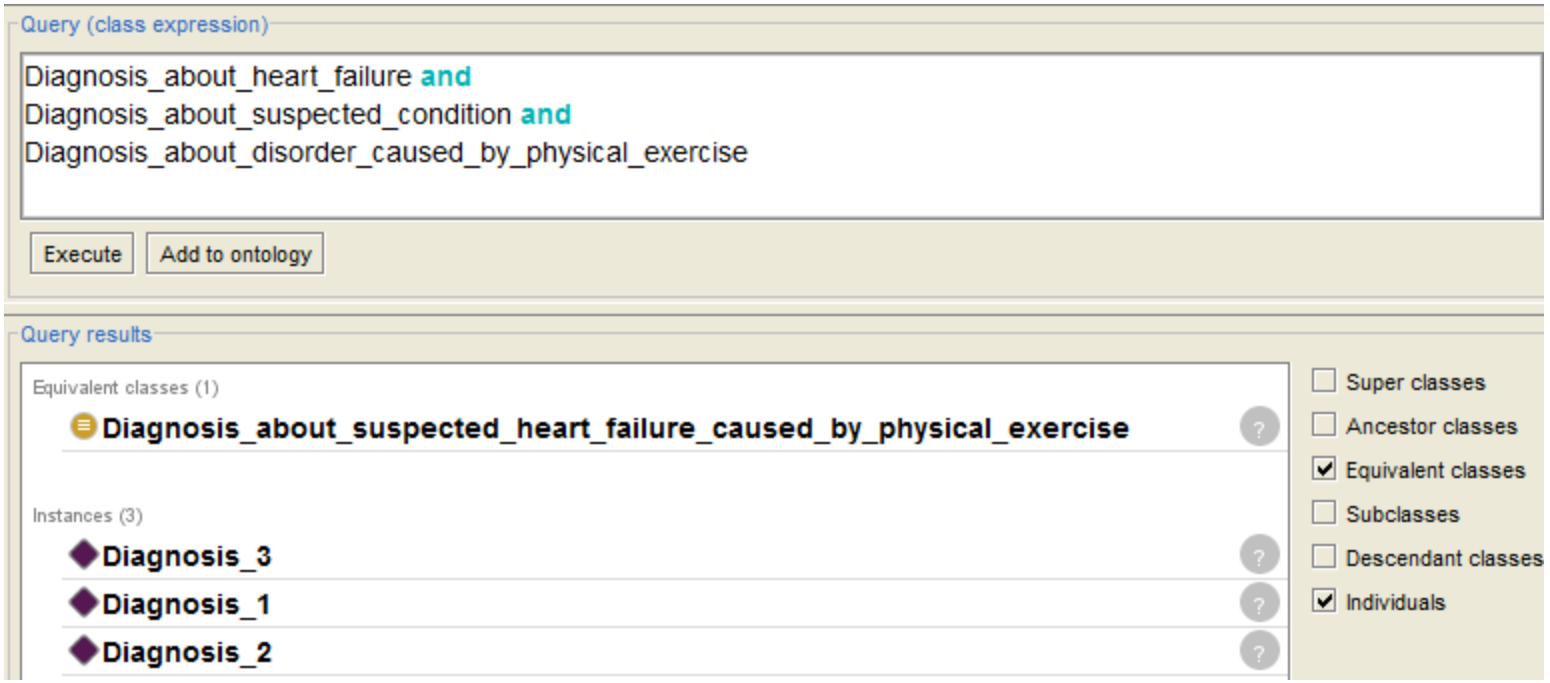
- **Diagnosis**
- **Diagnosis_about_condition**
- **Diagnosis_about_disorder_caused_by_physical_exercise**
- **Diagnosis_about_heart_disorder**
- **Diagnosis_about_heart_failure**
- **Diagnosis_about_heart_failure_caused_by_physical_exercise**
- **Diagnosis_about_organ_failure**
- **Diagnosis_about_suspected_condition**
- **Diagnosis_about_suspected_organ_failure**
- **InformationArtefact**

Instances (3)

- **Diagnosis_3**
- **Diagnosis_1**
- **Diagnosis_2**

All three information instances found

DL querying (II)

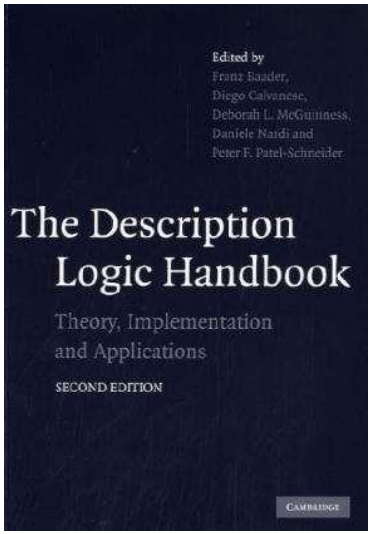


Ontologies for KR ?

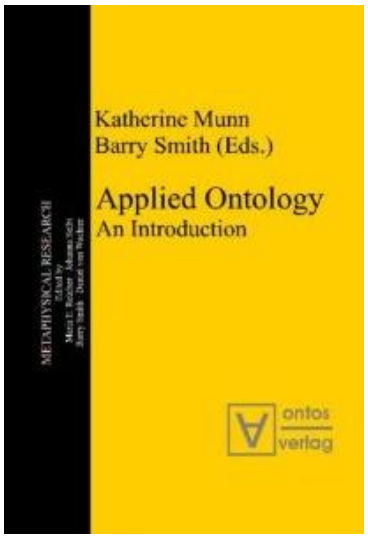
- Medical Ontologies provide coverage for most concepts referred to by medical KR approaches
- Medical terms and ontological knowledge should not be re-invented by KR systems
- Caveats: persisting content errors in biomedical ontologies
- Facts from the EHR encoded by ontologies (or terminologies) must not be interpreted out of context.
- Example on ontology-based annotation of information models a possible model for the interplay between KR and ontology artefacts?

Literature


Ludger Jansen,
Barry Smith (Hrsg.)
BIOMEDIZINISCHE ONTOLOGIE
Wissen strukturieren für den Informatik-Einsatz


Edited by
Franz Baader,
Diego Calvanese,
Deborah L. McGuinness,
Daniele Nardi and
Peter F. Patel-Schneider
The Description Logic Handbook
Theory, Implementation and Applications
SECOND EDITION
CAMBRIDGE


S. STAAB
R. STUDER
Editors
International Handbooks on Information Systems
HANDBOOK ON ONTOLOGIES
Springer


METAPHYSICAL RESEARCH
Katherine Munn
Barry Smith (Eds.)
Applied Ontology
An Introduction
ontos verlag


applied ontology


FORMAL ONTOLOGY IN INFORMATION SYSTEMS
IOS


JOURNAL OF BIOMEDICAL SEMANTICS


JOURNAL OF
Web Semantics
SCIENCE, SERVICES & AGENTS ON THE WORLD WIDE WEB

www

- Description Logics: http://dl.kr.org/
- Protégé: http://protege.stanford.edu/
- Bioontologiea: http://www.bioontology.ch/
- Buffalo Ontology Site: http://ontology.buffalo.edu/smith/
- OBO Foundry: http://obofoundry.org/
- Bioportal: http://bioportal.bioontology.org/
- SNOMED CT: http://www.ihtsdo.org/snomed-ct/
  http://terminology.vetmed.vt.edu/sct/menu.cfm