VIP

Virtual Imaging Platform

Stefan Schulz

Medical University of Graz (Austria)

purl.org/steschu

Does Medical Image Simulation Require Formal Ontologies?

VIP – Virtual Imaging Workshop

Lyon, France, Dec 14th, 2012

What are (formal) Ontologies?

What are (formal) Ontologies?

- Computer science view
  - Ontologies are purpose-oriented formal models of meaning (conceptualizations)
- Cognitive / linguistic view
  - Ontologies are concept systems or systems of semantic reference (no clear distinction from thesauri)
  - Also adopted by parts of the Semantic Web community
- Philosophy view (scientific realism)
  - Ontology is the study of what there is
  - Formal Ontologies give precise mathematical formulations of the properties and relations of certain entities.

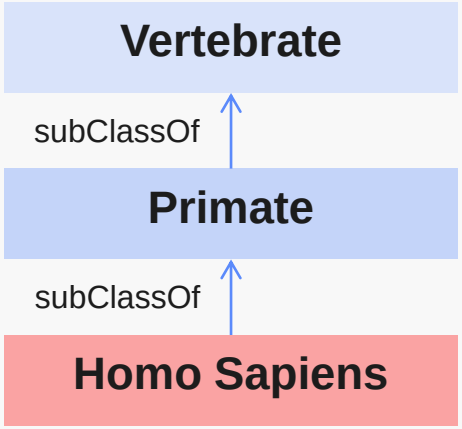
Quine O. On what there is. In: Gibson R. Quintessence - Basic Readings from the Philosophy of W. V. Quine. Cambridge: Belknap Press, Harvard University, 2004.
Schulz S, Stenzhorn H, Boeker M, Smith B: Strengths and limitations of formal ontologies in the biomedical domain. RECIIS - Electronic Journal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

Formal ontology in a nutshell

- Formal ontology = logic based ontology
- Description logics: subset of first order logic
- Common standard: OWL (Ontology Web Language)
- Ontologies are taxonomies of classes
- Ontologies can define classes in terms of (Aristotelian) definitions

Subclass (aka is-a):

A subClassOf *B*
iff
$\forall a{:}A(a) \rightarrow B(a)$



Primate subClassOf *Vertebrate*

Equivalence:

X equivalent to *Y* and some **r** *Z*
iff
$\forall x{:}X(x) \leftrightarrow Y(x) \wedge \exists z{:}C(z) \wedge \mathbf{r}(x,z)$

Vertebrate equivalentTo *Animal*
and **hasPart** some *Vertebra*

Taxonomy as Venn diagram

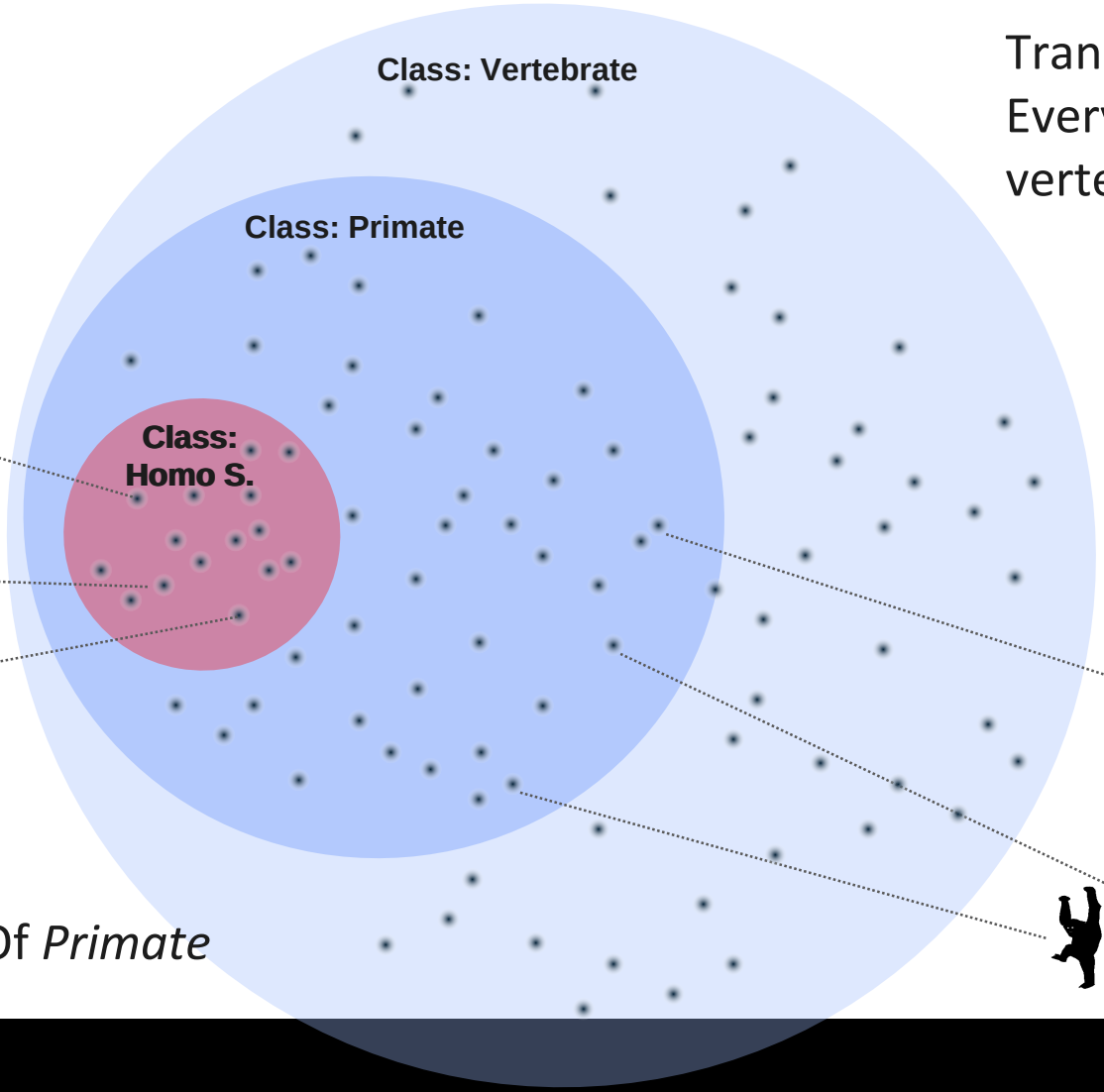
Every human is a primate, every primate is a vertebrate

Transitivity: Every human is a vertebrate

Tristan

Bernard

Stefan



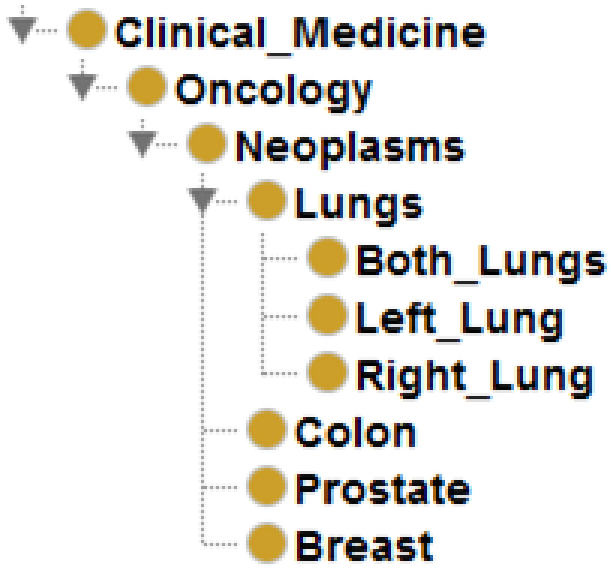
Washoe

Koko

Bobo

Human subClassOf *Primate*

Doing taxonomy right

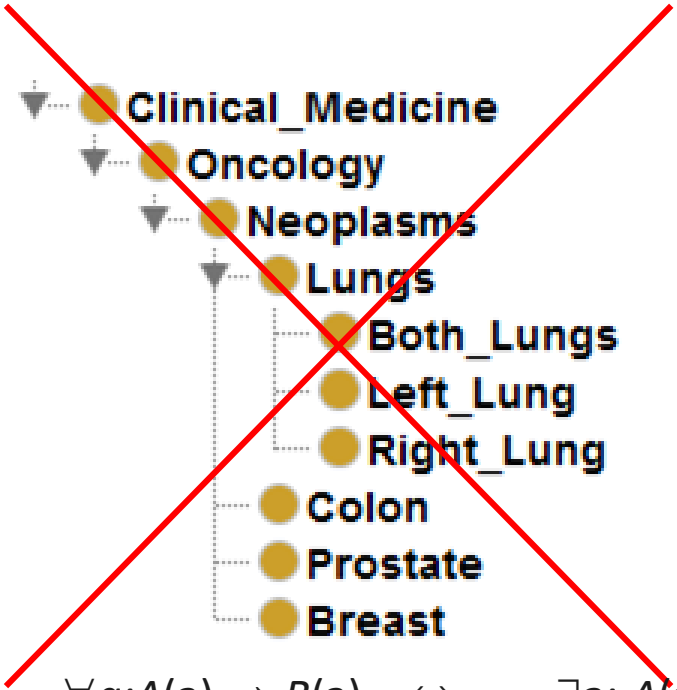


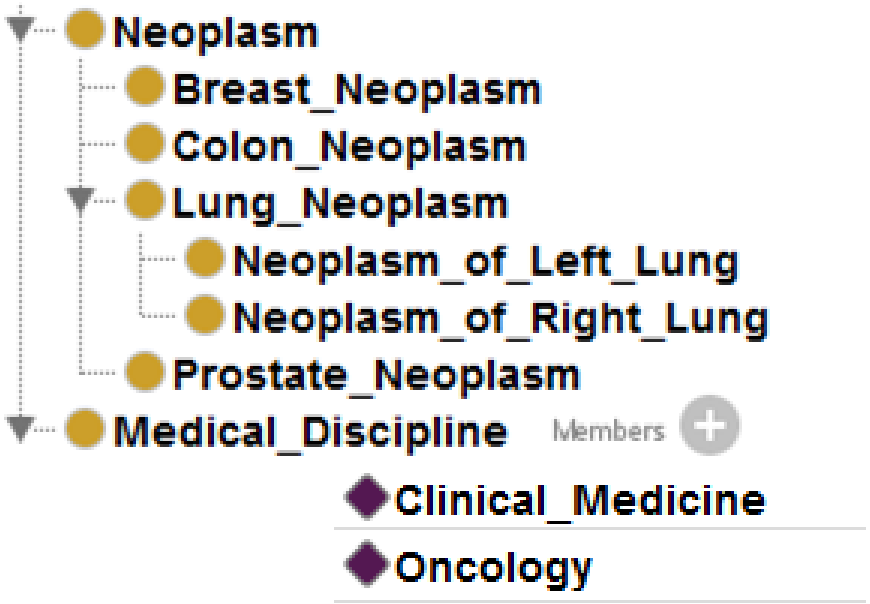
$\forall a{:}A(a) \rightarrow B(a) \quad \Leftrightarrow \quad \neg\exists a{:}\, A(a) \wedge \neg\, B(a)$

Test :

- there is no neoplasms that is not an oncology
- there is no prostate that is not a neoplasm
- there is no oncology that is not a clinical medicine

Doing taxonomy right





$\forall a{:}A(a) \rightarrow B(a) \quad \Leftrightarrow \quad \neg\exists a{:}\ A(a) \wedge \neg B(a)$

Test :

- oncology is an instance of a medical discipline
- there is no prostate neoplasm that is not a neoplasm

Labelling !

http://en.wikipedia.org/wiki/OntoClean

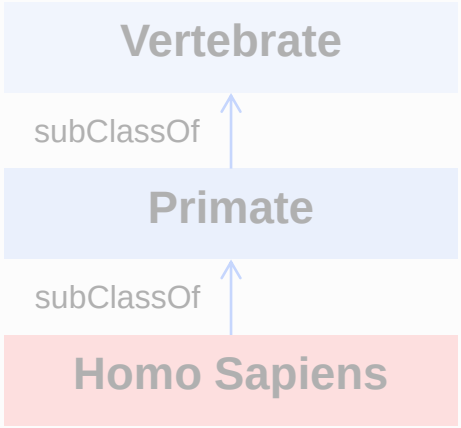
Schober D, Smith B, Lewis SE, Kusnierczyk W, Lomax J, Mungall C, Taylor CF, Rocca-Serra P, Sansone SA.Survey-based naming conventions for use in OBO Foundry ontology development. BMC Bioinformatics. 2009 Apr 27;10:125. doi: 10.1186/1471-2105-10-125.

Formal ontology in a nutshell

- Formal ontology = logic based ontology
- Description logics: subset of first order logic
- Common standard: OWL (Ontology Web Language)
- Ontologies are taxonomies of classes
- Ontologies can define classes in terms of (Aristotelian) definitions

Subclass (aka is-a):

A subClassOf *B*
iff
$\forall a{:}A(a) \rightarrow B(a)$



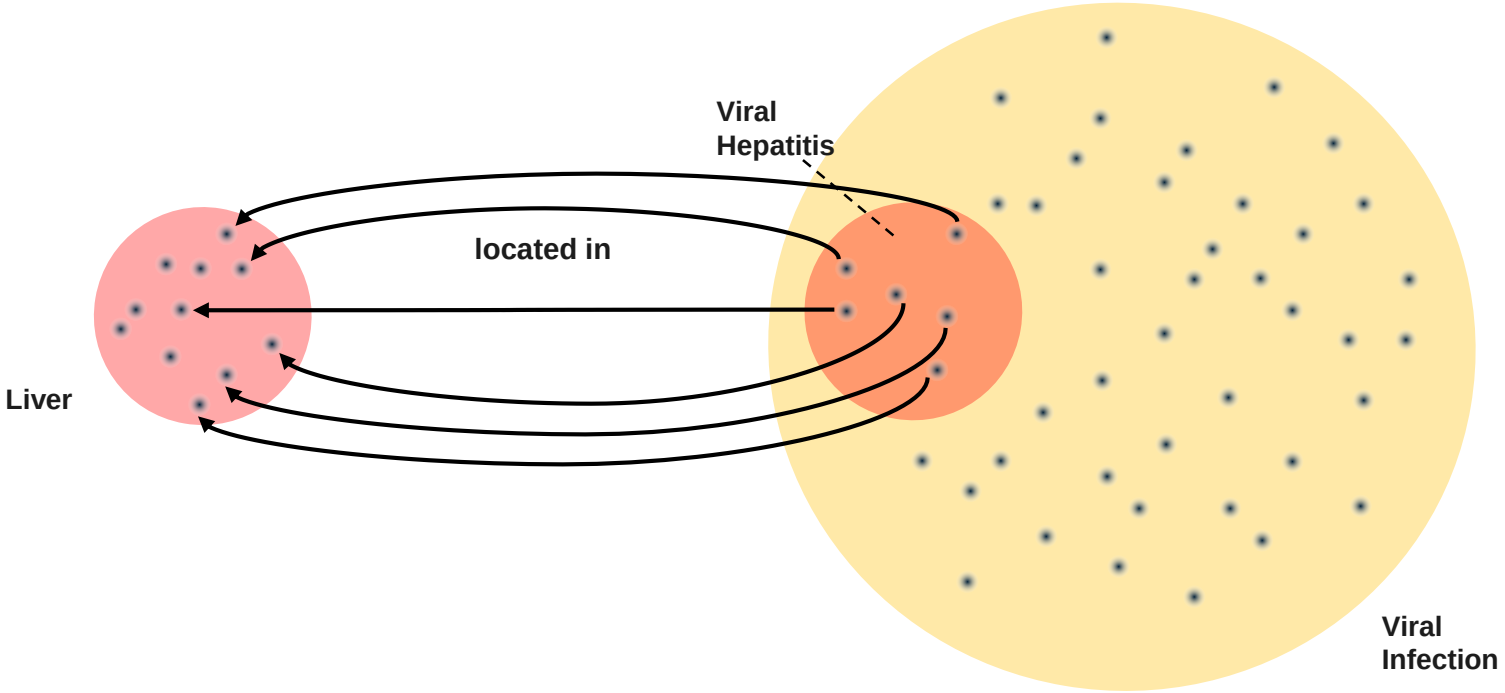
Primate subClassOf *Vertebrate*

Equivalence:

X equivalent to *Y* and some **r** *Z*
iff
$\forall x{:}X(x) \leftrightarrow Y(x) \wedge \exists z{:}\ C(z) \wedge \mathbf{r}(x,z)$

Vertebrate equivalentTo *Animal*
and **hasPart** some *Vertebra*

Aristotelian Definitions:
Genus proximum and *differentia specifica*

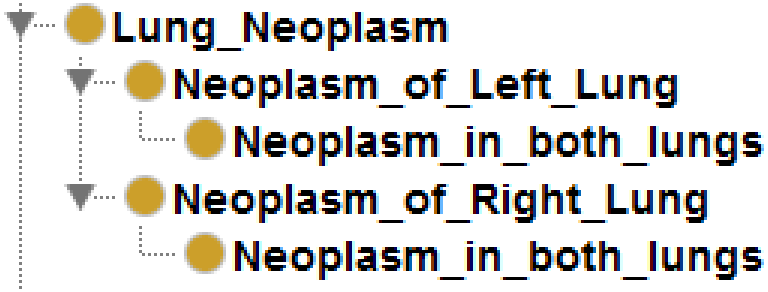


ViralHepatitis equivalentTo *ViralInfection* and **locatedIn** *Liver*

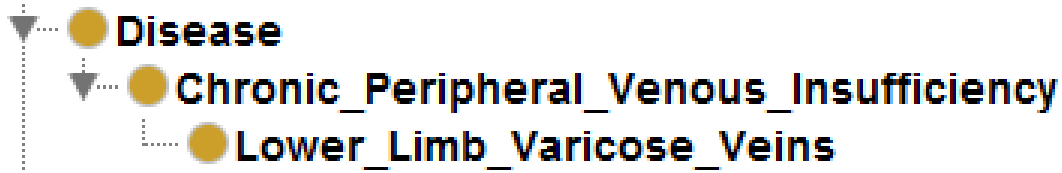
Test :

- There is no viral hepatitis that is not located in a liver
- There in no viral hepatitis that is not a viral infection

Ontological Commitment



- Which are the instances?
- Does the label tell us what is meant?
- Is there an implicit context?

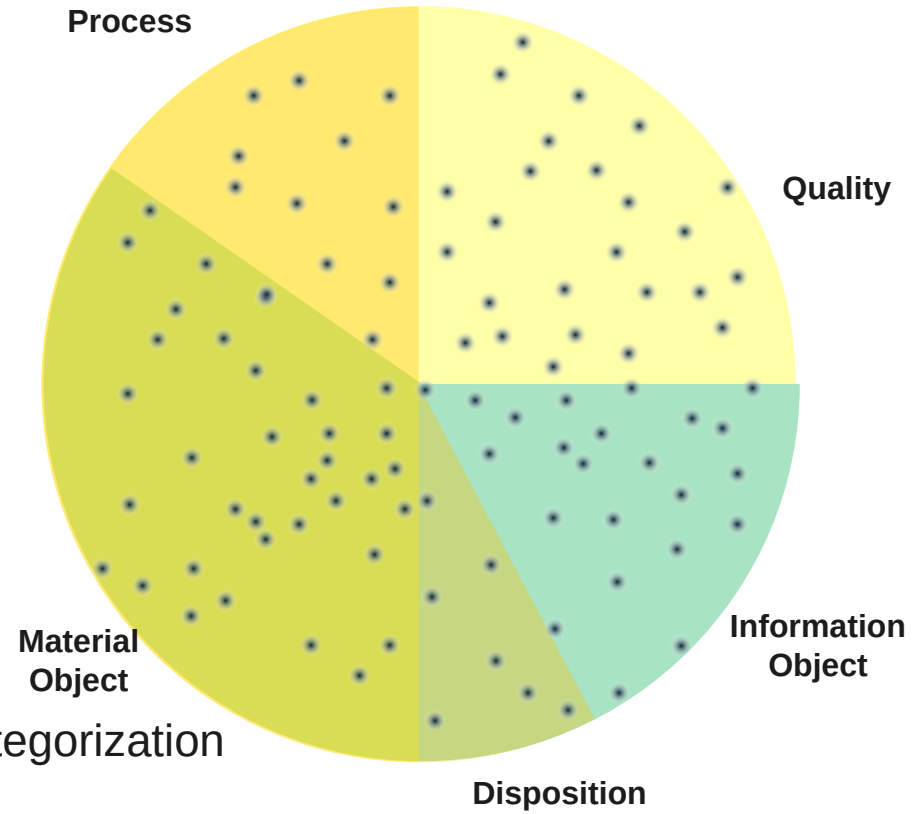


Test :

- There is no neoplasm in both lungs that is not a neoplasm in the left lung
- There is no varicose vein in the lower limb that is not a chronic peripheral venous insufficiency
- There is no patient with neoplasm in both lungs that is not a patient with the neoplasm in the left lung
- There is no patient with varicose lower limb veins that is not a patient with a chronic peripheral venous insufficiency

Upper level ontologies: partition of the domain into disjoint and exhaustive categories

- Particular
  - Quality
    - Benign
    - Malignant
  - Information_Object
    - Medical_Discipline
  - Disposition
    - Chronic_Peripheral_Venous_Insufficiency
  - Material_Object
    - Lower_Limb_Varicose_Veins
    - Neoplasm
      - Breast_Neoplasm
      - Colon_Neoplasm
      - Lung_Neoplasm
      - Prostate_Neoplasm
  - Process
    - Growth



- Upper level ontologies enforce a strict categorization
- Constraints on upper-level categories
- Upper level ontology for the biomedical domain **BioTop**

http://purl.org/biotop

How formal ontologies challenge human cognition

- Built around taxonomies of classes
  - ATTENTION: our intuitive way of hierarchically organize terms is not strictly taxonomic
- State what is true for all individual members of a class (instances of a type)
- Requires to distinguish between classes and individuals
  - ATTENTION: human language is often misleading, e.g. Lyon is a big city vs. The liver is a big organ
- Individuals commit to upper-level categories
  - ATTENTION: our thinking fuses mutually dependent entities that belong to different categories, e.g. Cancer (growth process vs. mass of malignant tissue)
- Upper level categories should be made explicit
  - Explicit upper level ontology – common understanding
  - Implicit upper level ontology of each of us – misunderstanding

What formal ontology is not

- Ontology ≠ Knowledge representation
  - "*There are very few interesting items of knowledge that are truly ontological in this strict sense*" (Alan Rector)
  - antinomy: *ὄντος* (being) vs. *ἐπιστήμη* (knowledge)
- Ontology is not appropriate for
  - Default knowledge
    - "The hand has 5 fingers" (unless otherwise stated)
  - Probabilistic knowledge
    - Mesothelioma is a rare cancer
  - Contingent knowledge
    - Aspirin prevents myocardial infarction
    - Jaundice is a typical symptom of hepatitis

Why formal ontology at all ??

- Formal definitions create maximum consensus on the meaning of terms
  - Ontologies as standards
  - Reusable terms and axioms
- Formal axioms encode statements about what is considered to be universally true in a domain
  - in contrast to knowledge proper
- Formal axioms permit logic-based reasoning
  - Consistency checking
  - New entailments
  - Equivalence of syntactically heterogeneous expressions can be computed: semantically interoperable systems

Ontologies in life sciences and health care

Ontologies in life sciences and health care

- Bottom-up ontology development:
  OBO (Open biomedical Ontologies) Foundry

- Top-down ontology development:
  SNOMED CT
  (Systematized Nomenclature of Medicine – Clinical terms)

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)

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

The OBO Foundry

- Collaborative bottom up initiative, driven by the success of the Gene Ontology
- Rooted in upper ontologies (BFO + RO)
- Goal of creating a suite of orthogonal interoperable reference ontologies in the biomedical domain
- Moving from semi-formal OBO syntax to OWL-DL
- Cross-ontology definitional axioms:
  - *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

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

SNOMED CT (Systematized Nomenclature of Medicine - Clinical Terms)

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.

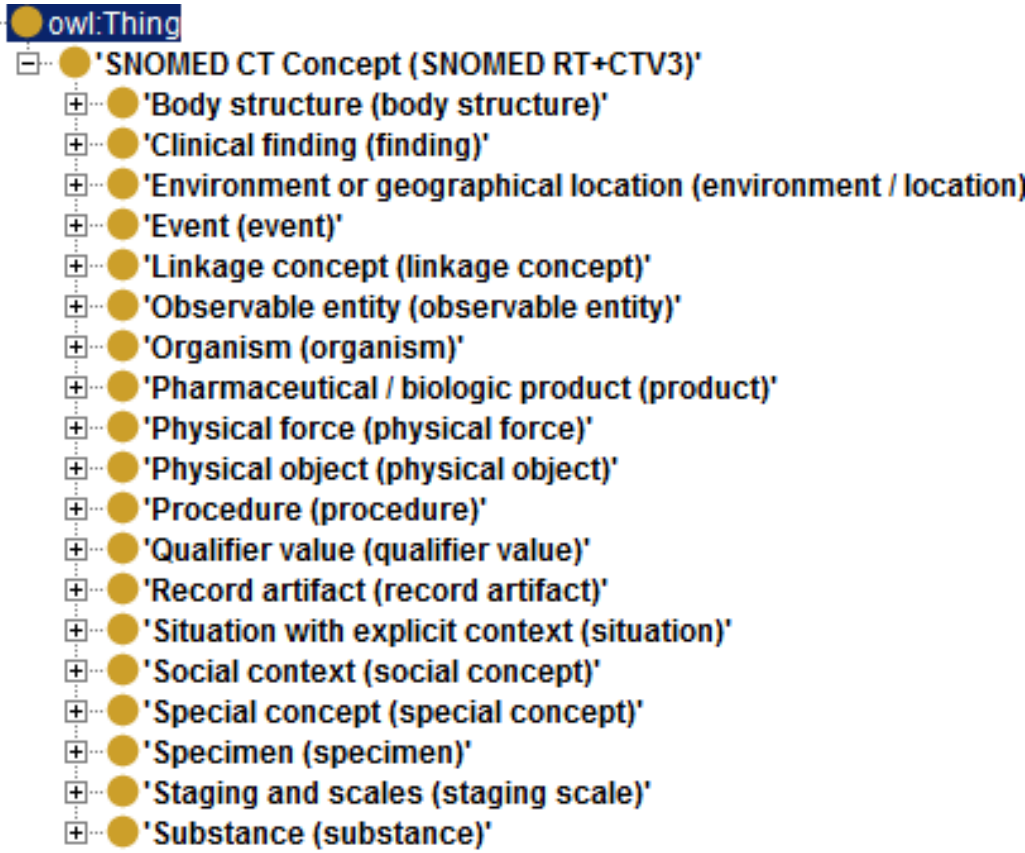
SNOMED CT intellectual property rights were transferred to the SNOMED SDO® in the formal creation of the IHTSDO.

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.

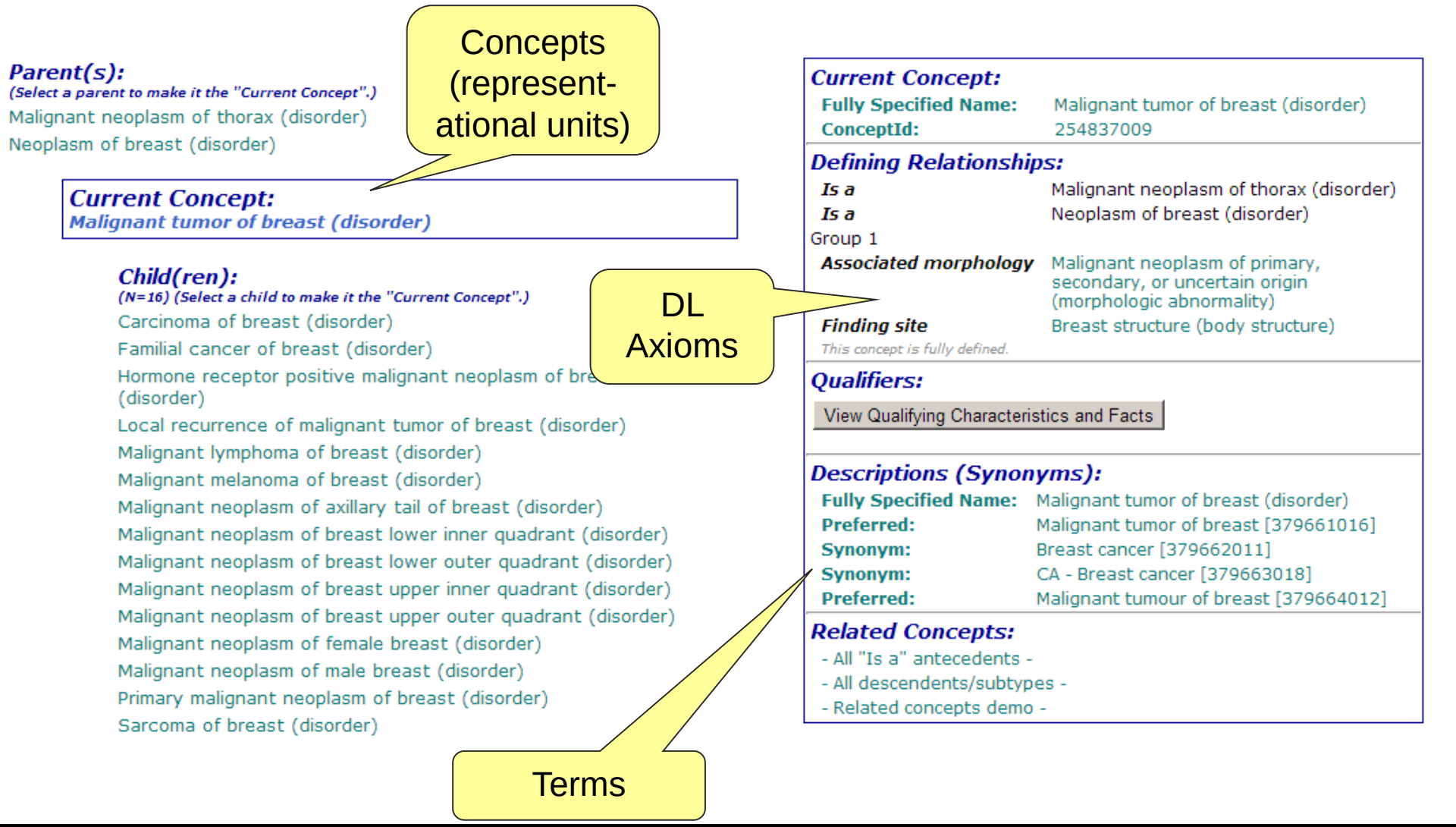
http://www.ihtsdo.org/

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
- Definitions with DL axioms
- Promoted as an international terminological standard"



SNOMED CT: Terminology + Ontology



http://viw2.vetmed.vt.edu/sct/menu.cfm

Bioportal – repository for biomedical ontologies

Term Search

Search for a term in multiple ontologies ?

myocardium

Search

advanced options

Myocardium - NCI Thesaurus
http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#Myocardium
The striated muscle tissue of the heart enveloped by the epicardium and the endocardium.
details - visualize - 2 more from this ontology

myocardium - Mouse adult gross anatomy
http://purl.obolibrary.org/obo/MA_0000164
details - visualize - 3 more from this ontology

myocardium - RadLex
http://purl.bioontology.org/ontology/RID/RID1398
details - visualize - 2 more from this ontology

Myocardium - Medical Subject Headings (MeSH)
http://purl.bioontology.org/ontology/MSH/D009206
The muscle tissue of the HEART. It is composed of striated, involuntary muscle cells (MYOCYTES, CARDIAC) connected to form the contractile pump to generate blood flow.
details - visualize - 1 more from this ontology

myocardium - Experimental Factor Ontology
http://www.ebi.ac.uk/efo/EFO_0000819
Muscle layer of organ which has as its parts the myocardium proper and the conducting system of the heart.
details - visualize - 2 more from this ontology

Myocardium - Logical Observation Identifier Names and Codes
http://purl.bioontology.org/ontology/LNC/LP101987-8
details - visualize - 6 more from this ontology

Myocardium - Foundational Model of Anatomy
http://sig.uw.edu/fma#Myocardium
Muscle layer of organ which has as its parts the myocardium proper and the conducting system of the heart.
details - visualize - 1 more from this ontology

Myocardium - Cell line ontology
http://purl.obolibrary.org/obo/FMA_9462
details - visualize - 6 more from this ontology

Myocardium - Read Codes, Clinical Terms Version 3 (CTV3)
http://purl.bioontology.org/ontology/RCD/XC0Is
details - visualize - 5 more from this ontology

myocardium - Zebrafish anatomy and development
http://purl.obolibrary.org/obo/ZFA_0001319
details - visualize - 2 more from this ontology

myocardium - BRENDA tissue / enzyme source
http://purl.obolibrary.org/obo/BTO_0000901
The middle and thickest layer of the heart wall, composed of cardiac muscle.
details - visualize

myocardium - Teleost Anatomy Ontology
http://purl.obolibrary.org/obo/TAO_0001319
details - visualize - 2 more from this ontology

myocardium - CRISP Thesaurus, 2006
http://purl.bioontology.org/ontology/CSP/1390-2774
muscle tissue of the heart composed of striated, involuntary muscle known as cardiac muscle.
details - visualize - 1 more from this ontology

myocardium - eVOC (Expressed Sequence Annotation for Humans)
http://purl.obolibrary.org/obo/EV_0100022
The middle layer of the heart wall that consists of cardiac muscle.
details - visualize - 2 more from this ontology

Myocardium - Cardiac Electrophysiology Ontology
http://purl.org/obo/owl/FMA#FMA_9462
details - visualize - 1 more from this ontology

myocardium - Minimal anatomical terminology
http://purl.obolibrary.org/obo/MAT_0000453
details - visualize

Myocardium - BioModels Ontology
http://purl.org/obo/owlapi/fma#FMA_9462
details - visualize - 1 more from this ontology

Submit your ontology to
http://bioportal.bioontology.org

Formal ontologies and beyond…

http://bioportal.bioontology.org

Alternative to formal ontologies: INFORMAL terminologies / thesauri

Alternative to formal ontologies: INFORMAL terminologies / thesauri

- Group 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

Freitas F, Schulz S, Moraes E: Survey of current terminologies and ontologies in biology and medicine. RECIIS - Electronic Journal in Communication,Information and Innovation in Health, 2009; 3 (1): 7-18:http://dx.doi.org/10.3395/reciis.v3i1.239en

MeSH - Medical Subject Headings

- 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] +

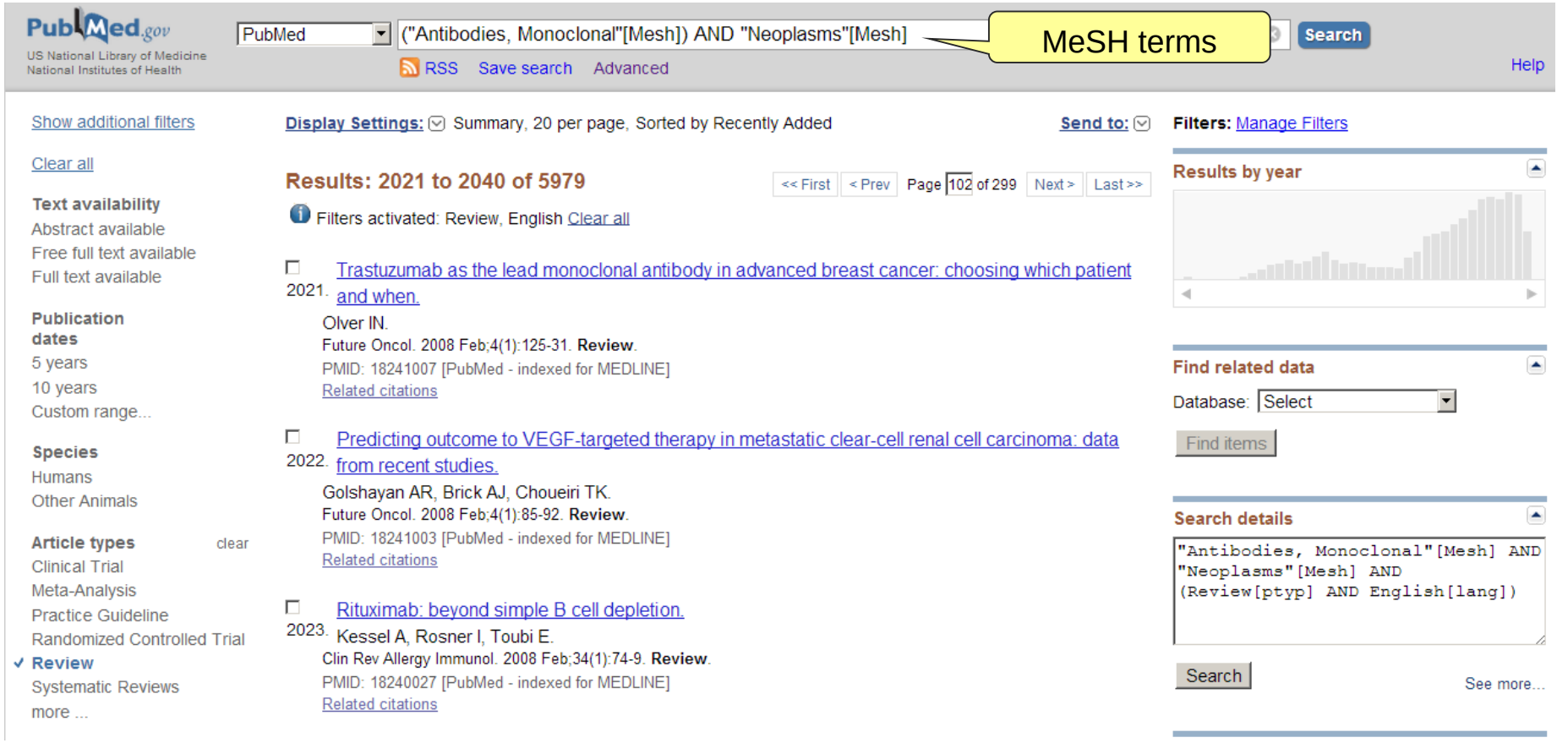
- Skin and Connective Tissue Diseases [C17]
  - Skin Diseases [C17.800]
    - Breast Diseases [C17.800.090]
      - Breast Cyst [C17.800.090.249]
      - Breast Neoplasms [C17.800.090.500]
        - Breast Neoplasms, Male [C17.800.090.500.260]
        - Carcinoma, Ductal, Breast [C17.800.090.500.390]
        - 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] +

http://www.nlm.nih.gov/mesh/

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 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 R
Entry Version	BREAST NEOPL
Date of Entry	19990101
Unique ID	D001943

Example: Medical Subject Headings (MeSH)

Literature search: monoclonal antibodies and cancer therapy

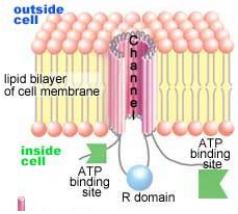


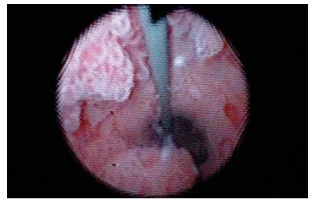
"Nontologies"

- Use OWL syntax, which should not be interpreted according to description logics semantics
- Formal reasoning would lead to incorrect entailments
- Examples: NCI thesaurus, Radlex
- Many other ontologies contain problematic axioms that contradict the intended meaning
- Example (NCI thesaurus):

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*

Ureter_Small_Cell_Carcinoma subclassOf
Disease_May_Have_Finding some *Pain*





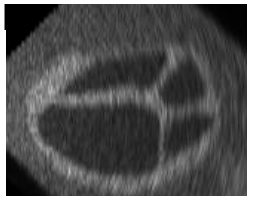
Schulz S, Schober S, Tudose I, Stenzhorn H: The Pitfalls of Thesaurus Ontologization – the Case of the NCI Thesaurus. AMIA Annu Symp Proc, 2010: 727-731

(N)ontologies of interest for imaging

- Radlex
  - 24800 classes covering anatomy, procedures, diseases, substances, devices, relevant for radiologic imaging
  - all classes are also individuals (punning)
  - Relations ('is a', 'part of') asserted are at the level of individuals
  - On classes no other axioms than subclass axioms
- Foundational model of anatomy
  - Complete model: Protégé Frames (no formal semantics)
    Parts of it available as OWL
  - All assertions at class level
  - Logical entailments only true for "canonical" anatomy
- SNOMED CT

Challenges of "correct" ontology for image representation including simulation

- Same terms (e.g. "cardiac motion") can be used for different things
  1. A real cardiac motion in a patient (process)
  2. Part of an image (information entity) that represents a real cardiac motion
  3. A simulation artifact (information entity), which does not refer to any specific cardiac motion
  4. The "concept" cardiac motion (cognitive entity)



- If you prefer 4. or if the distinction between 1. – 4. does not matter, then you shouldn't use formal ontologies

Where ontologies or thesauri are sufficient

- Provision of controlled terms
  - Good text definitions should be available
- Hierarchy expansion for retrieval
  - hierarchical links at the level of broader term / narrower terms
  - is-a → is narrower than
  - part-of → is narrower than
- "Hand-crafted" inference rules
  - no use of description logics classifiers
- Possible standard SKOS

Own experiences with ontologies in large projects

- EU funded projects with multiple partners
  1. @neurIST: Data integration (clinical, genomic, simulation) on cerebral aneurysms
  2. DebugIT: Decision support system for infectious diseases
  3. SemanticHealthNet: Semantic interoperability between heterogeneous semantic representations in the EHR
- Experiences:
  - in 1. and 2. much effort put in formal ontology
  - Mostly used as a controlled vocabulary (1.)
  - DL reasoning only for computing inferred ontology, which then used with production rules
  - 3. Formal foundation seems fundamental to reach the interoperability goal. However, intellectual input considerable and scalability still open

Current state of the art of Applied Ontology as a discipline

Current state of the art of Applied Ontology as a discipline

- Applied Ontology – still emerging discipline
- Prevalence of makeshift ontology artifacts
- Ontology engineering required to be more principled
- Necessary resources
  - Standards (Semantic Web – OWL)
  - Good practice guidelines (e.g. **GoodOD** Guideline)
  - Quality management
  - Best-of-breed examples
  - Industry-standard tools
    - Editors
    - Reasoners

Current state of the art of Applied Ontology as a discipline

Guideline on Developing Good Ontologies in the Biomedical Domain with Description Logics

URL: http://www.purl.org/goodod/guideline

Version 1.0
December 2012

Send feedback to:
martin.boeker@uniklinik-freiburg.de
ludger.jansen@uni-rostock.de

Schulz S[1,3], Seddig-Raufie D[1], Grewe N[2], Röhl J[2],
Schober D[1], Boeker M[1], Jansen L[2]

[1]: Institute of Medical Biometry and Medical Informatics,
University Medical Center Freiburg
[2]: Institute of Philosophy, University of Rostock
[3]: Department of Medical Informatics, University of Graz

11th December 2012

http://www.iph.uni-rostock.de/GoodOD-Guideline.1299.0.html

Current state of the art of Applied Ontology as a discipline

- Applied Ontology – still emerging discipline
- Prevalence of makeshift ontology artifacts
- Ontology engineering required to be more principled
- Necessary resources
  - Standards (Semantic Web – OWL)
  - Good practice guidelines (e.g. **GoodOD** Guideline)
  - Quality management
  - Best-of-breed examples
  - Industry-standard tools
    - Editors
    - Reasoners

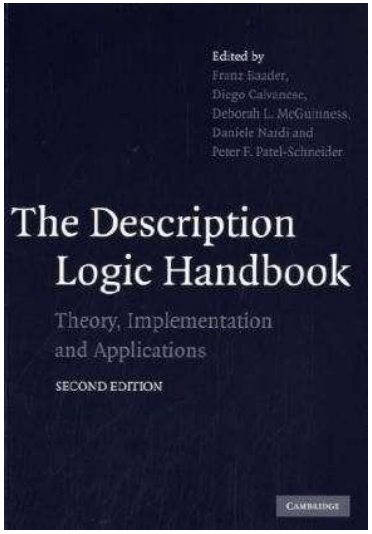
Decision on using formal ontology in life science research projects

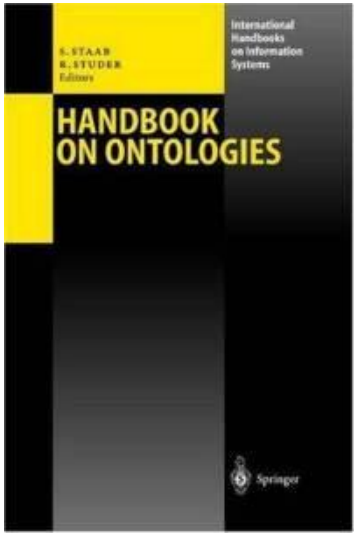
- Functional requirements
  - Controlled terminology
  - Query expansion for retrieval
  - Precise definitions of terms
  - Precise classification of domain entities
  - Reasoning to establish semantic equivalence
  - Representation of contingent knowledge
  - Default reasoning
  - Probabilistic reasoning

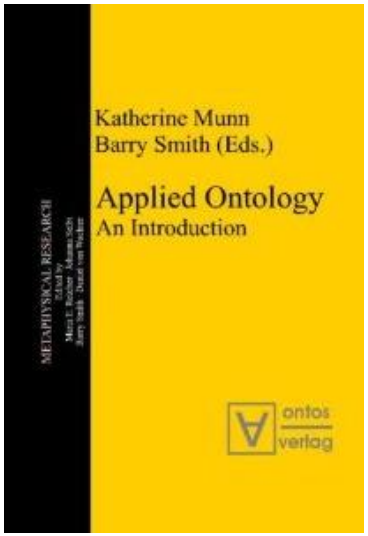
Decision matrix
Thesauri / Ontologies / KR formalism

	thesauri	ontology	KR
Controlled domain language	x		
Query expansion for retrieval	x	x	
Precise definitions of terms		x	
Precise classification of domain entities		x	
Reasoning to establish semantic equivalence		x	
Representation of contingent knowledge			x
Non-monotonic reasoning			x
Probabilistic reasoning			x

Further readings







Ontology on the Web

- Description Logics: http://dl.kr.org/
- Protégé: http://protege.stanford.edu/
- Bioontology: 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
- CO-ODE (Pizza ontology): http://www.co-ode.org/
- GoodOD Guideline: http://www.iph.uni-rostock.de/GoodOD-Guideline.1299.0.html