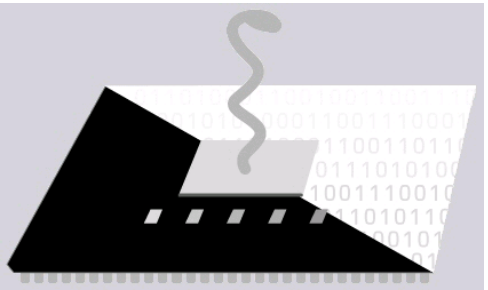
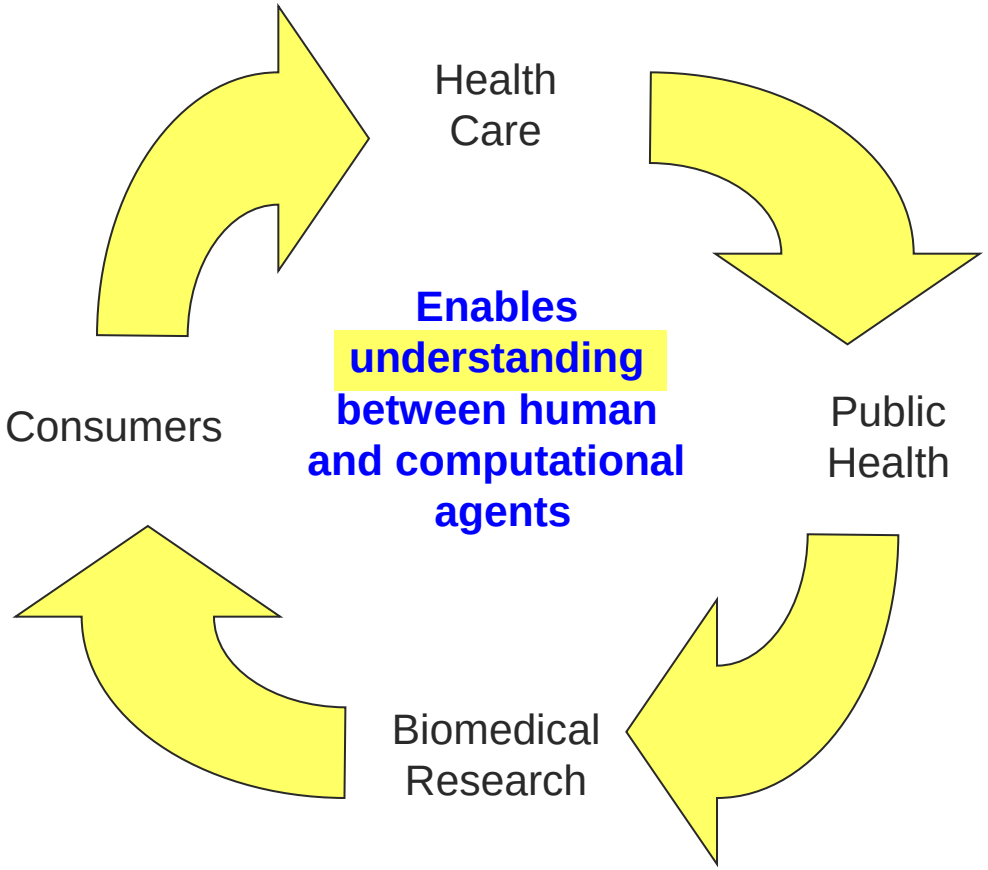
Ten Theses on Biomedical Ontologies and Terminology Systems

Stefan Schulz, Holger Stenzhorn
Freiburg University Hospital
Department of Medical Informatics
Medical Language and Ontology Group (MediLOG)



ICMCC
the international council on medical & care compunetics

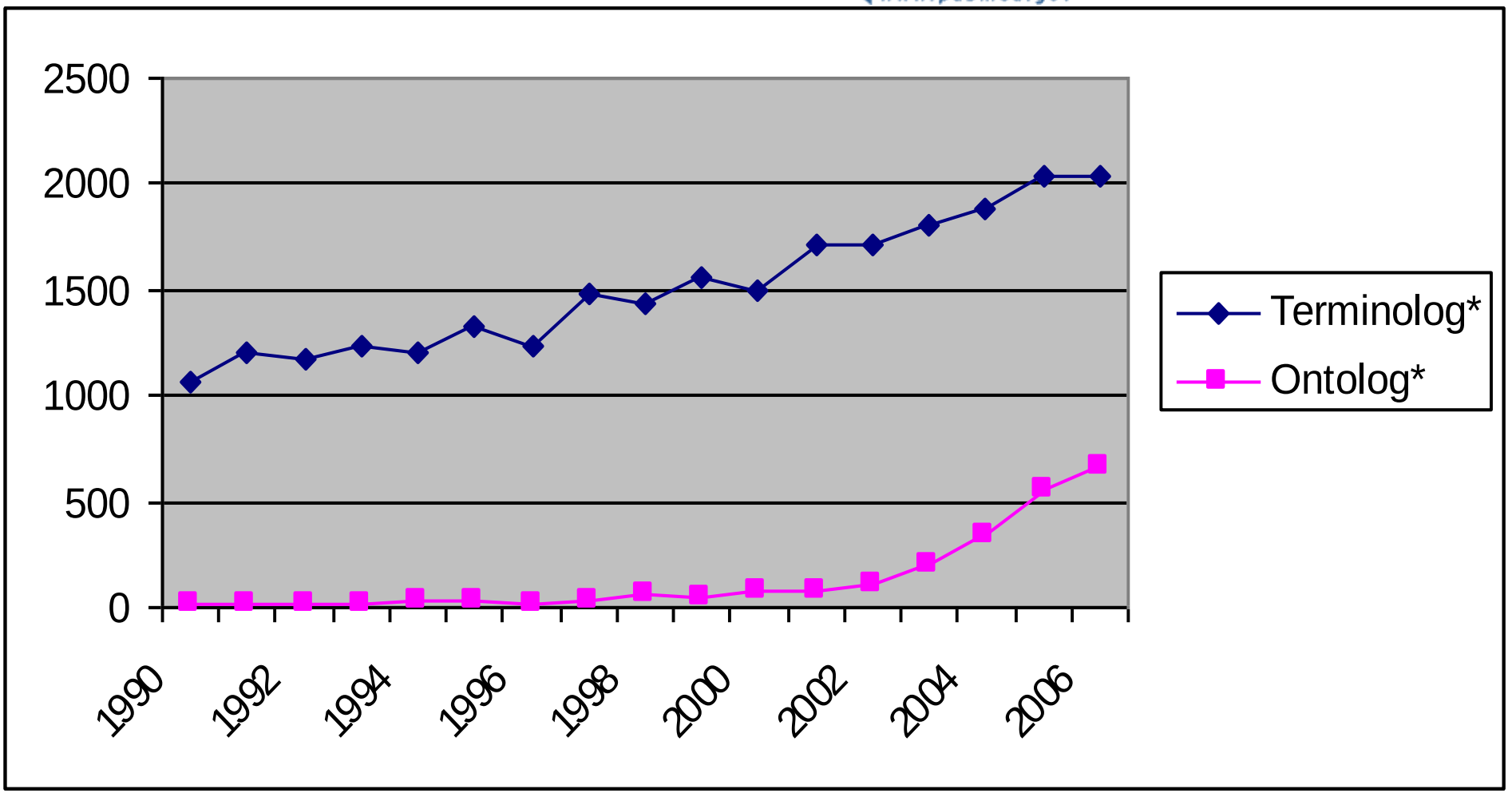
Semantic Interoperability



Common language: Ontologies and Terminology Systems

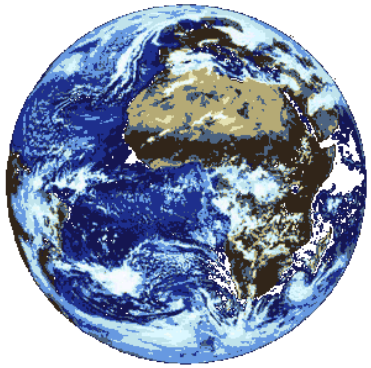
Literature on Biomedical Terminologies and Ontologies

PubMed www.pubmed.gov
A service of the National Library of Medicine and the National Institutes of Health



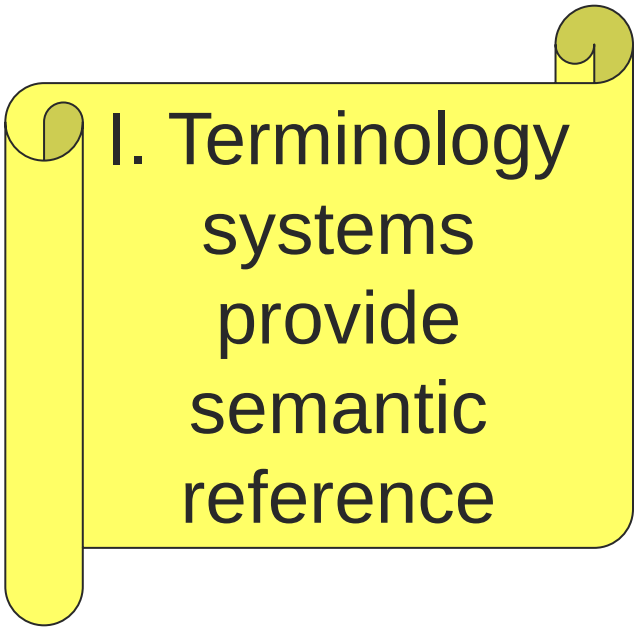
Definitions

- Terminology System
  - Set of terms representing the system of concepts of a particular subject field. (ISO 1087)
- Ontology / Formal Ontologies
  - Ontology is the study of w
  - Formal ontologies are theor
    give precise mathematical fo
    properties and relations of ce
    (Stanford Encyclopedia of Ph

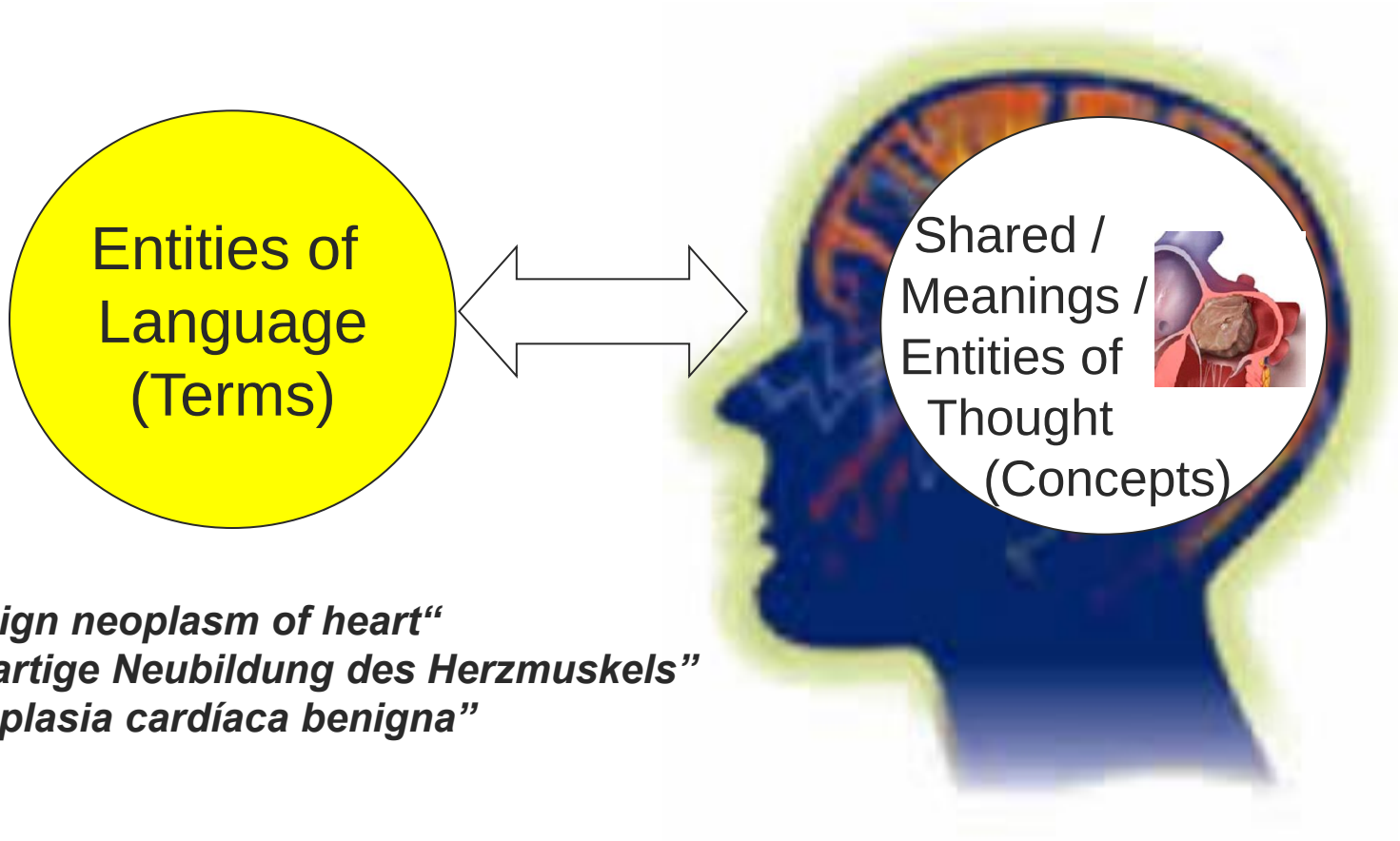


Purpose of this talk

- Make clear the differences between Terminology Systems and Ontologies
- Propose ten theses to characterize the nature, the rationale, and the limitations of (clinical, biomedical) ontologies
- Initiate discussion about
  - Preference for terminologies or (formal) ontologies in practical applications
  - How to “ontologize” existing terminologies
  - Co-existence between terminologies and ontologies


I. Terminology
systems
provide
semantic
reference

Semantic Reference



„benign neoplasm of heart“
„gutartige Neubildung des Herzmuskels”
“neoplasia cardíaca benigna”

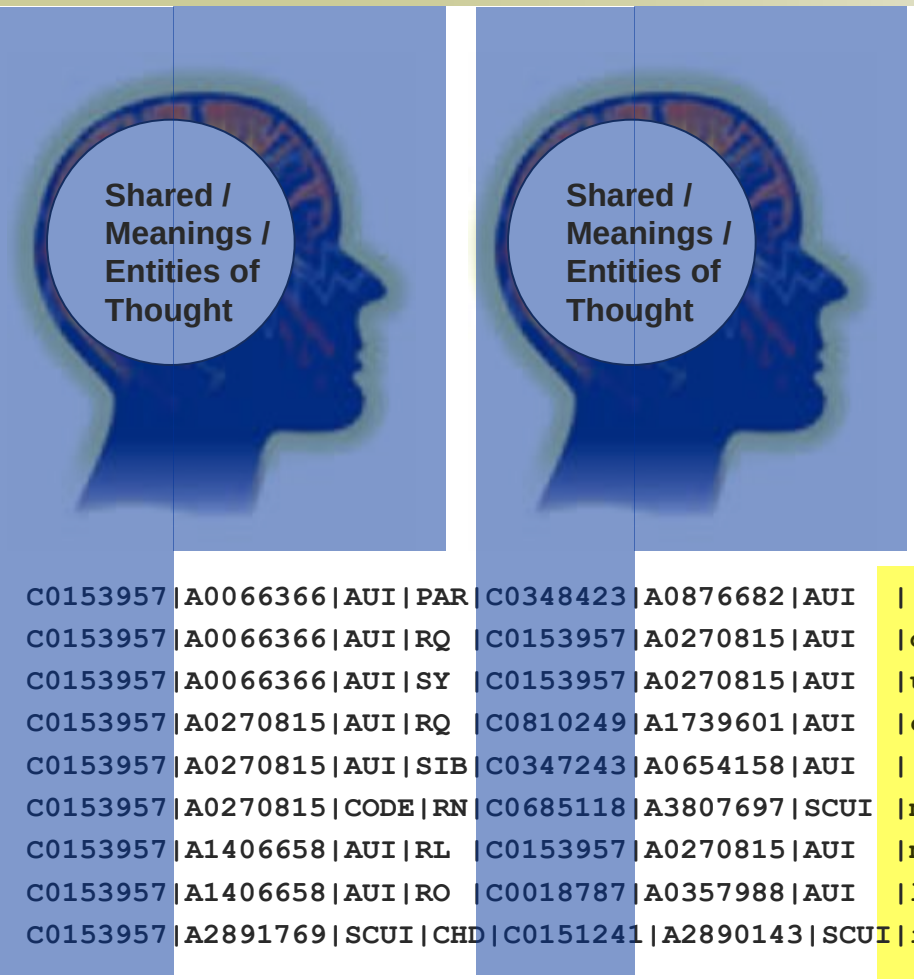
Example: UMLS (mrconso table)

Entities of Language (Terms)

```
C0153957|ENG|P|L0180790|PF|S1084242|Y|A1141630||||MTH|PN|U001287|benign neoplasm of heart|0|N||
C0153957|ENG|P|L0180790|VC|S0245316|N|A0270815||||ICD9CM|PT| 212.7|Benign neoplasm of heart|0|N||
C0153957|ENG|P|L0180790|VC|S0245316|N|A0270817||||RCD|SY|B727.| Benign neoplasm of heart|3|N||
C0153957|ENG|P|L0180790|VO|S1446737|Y|A1406658||||SNMI|PT|    D3-F0100|Benign neoplasm of heart, NOS|3|N||
C0153957|ENG|S|L0524277|PF|S0599118|N|A0654589||||RCDAE|PT|B727.|Benign tumor of heart|3|N||
C0153957|ENG|S|L0524277|VO|S0599510|N|A0654975||||RCD|PT|B727.| Benign tumour of heart|3|N||
C0153957|ENG|S|L0018787|PF|S0047194|Y|A0066366||||ICD10|PS|D15.1|Heart|3|Y||
C0153957|ENG|S|L0018787|VO|S0900815|Y|A0957792||||MTH|MM|U003158|Heart <3>|0|Y||
C0153957|ENG|S|L1371329|PF|S1624801|N|A1583056|||10004245|MDR|LT|10004245|Benign cardiac neoplasm|3|N||
C0153957|GER|P|L1258174|PF|S1500120|Y|A1450314||||DMDICD10|PT| D15.1|Gutartige Neubildung: Herz|1|N||
C0153957|SPA|P|L2354284|PF|S2790139|N|A2809706||||MDRSPA|LT| 10004245|Neoplasia cardiaca benigna|3|N||
```

Unified Medical Language System, Bethesda, MD: National Library of Medicine, 2007: http://umlsinfo.nlm.nih.gov/

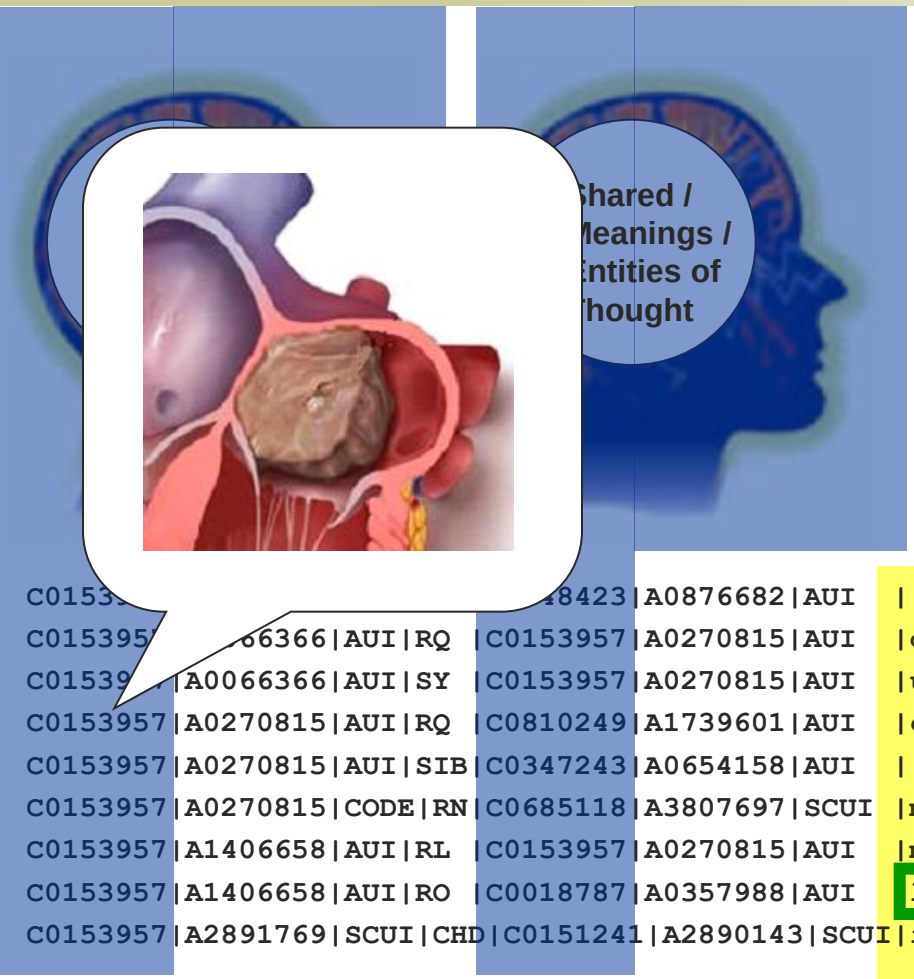
Example: UMLS (mrrel table)

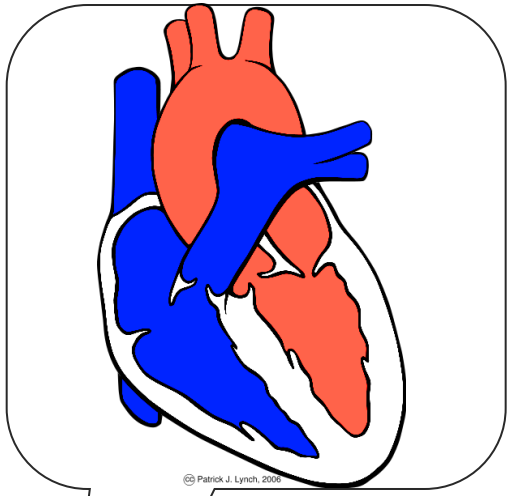



```
C0153957|A0066366|AUI|PAR|C0348423|A0876682|AUI  |                    |R06101405||ICD10|ICD10|||N||
C0153957|A0066366|AUI|RQ |C0153957|A0270815|AUI  |default_mapped_ from|R03575929||NCISEER|NCISEER|||N||
C0153957|A0066366|AUI|SY |C0153957|A0270815|AUI  |uniquely_mapped_ to |R03581228||NCISEER|NCISEER|||N||
C0153957|A0270815|AUI|RQ |C0810249|A1739601|AUI  |classifies          | R00860638||CCS|CCS|||N||
C0153957|A0270815|AUI|SIB|C0347243|A0654158|AUI  |                    |R06390094          || ICD9CM|ICD9CM|
C0153957|A0270815|CODE|RN|C0685118|A3807697|SCUI |mapped_to           | R15864842||SNOMEDCT|SNOMEDCT||Y|N||
C0153957|A1406658|AUI|RL |C0153957|A0270815|AUI  |mapped_from         | R04145423||SNMI|SNMI|||N||
C0153957|A1406658|AUI|RO |C0018787|A0357988|AUI  |location_of         | R04309461||SNMI|SNMI|||N||
C0153957|A2891769|SCUI|CHD|C0151241|A2890143|SCUI|isa                 |R19841220|47189027|SNOMEDCT|SNOMEDCT
```

Semantic relations

Example: UMLS






```
C0153[illegible]|[illegible]8423|A0876682|AUI  |                    |R061[illegible]|ICD10|ICD10|||N||
C015395[illegible]66366|AUI|RQ  |C0153957|A0270815|AUI  |default_mapped_ from|R035[illegible]||NCISEER|NCISEER|||N||
C01539[illegible]|A0066366|AUI|SY  |C0153957|A0270815|AUI  |uniquely_mapped_ to |R035[illegible]8||NCISEER|NCISEER|||N||
C0153957|A0270815|AUI|RQ  |C0810249|A1739601|AUI  |classifies          | R0[illegible]0638||CCS|CCS|||N||
C0153957|A0270815|AUI|SIB|C0347243|A0654158|AUI  |                    |R06[illegible]0094          || ICD9CM|ICD9CM|
C0153957|A0270815|CODE|RN|C0685118|A3807697|SCUI |mapped_to           | R1[illegible]864842||SNOMEDCT|SNOMEDCT||Y|N||
C0153957|A1406658|AUI|RL  |C0153957|A0270815|AUI  |mapped_from         | R[illegible]4145423||SNMI|SNMI|||N||
C0153957|A1406658|AUI|RO  |C0018787|A0357988|AUI  |location_of         | R04309461||SNMI|SNMI|||N||
C0153957|A2891769|SCUI|CHD|C0151241|A2890143|SCUI|isa                 |R19841220|47189027|SNOMEDCT|SNOMEDCT
```

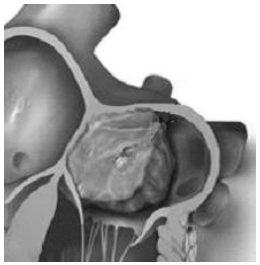
INFORMAL **Semantic relations**

II. Ontologies are taxonomies of semantic types that support the organization of domain entities

Organizing Entities

Entity Types

The type “benign neoplasm of heart”

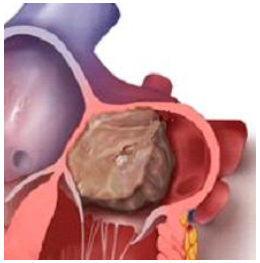


Entities of Language

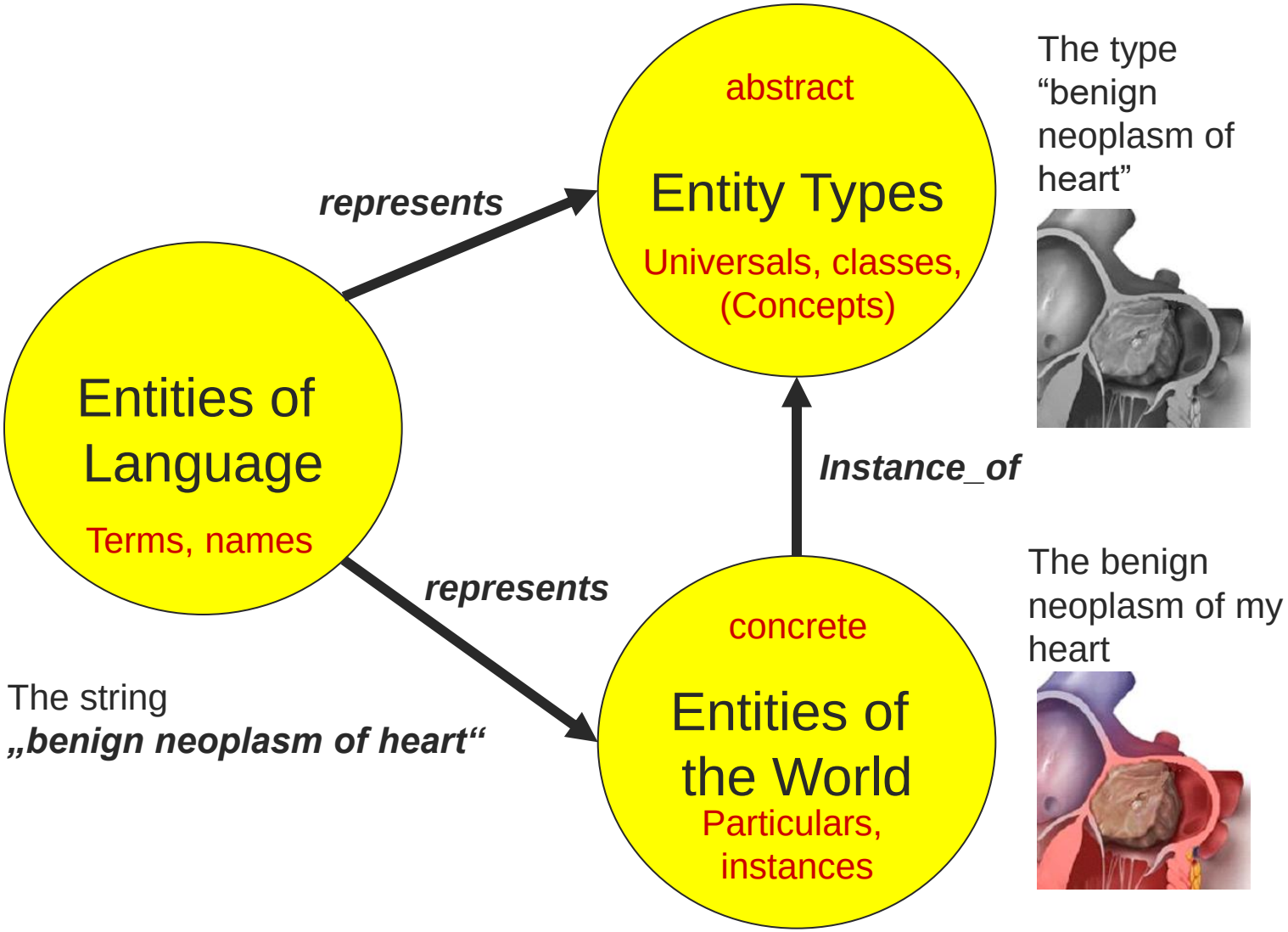
The string
„benign neoplasm of heart“

Entities of the World

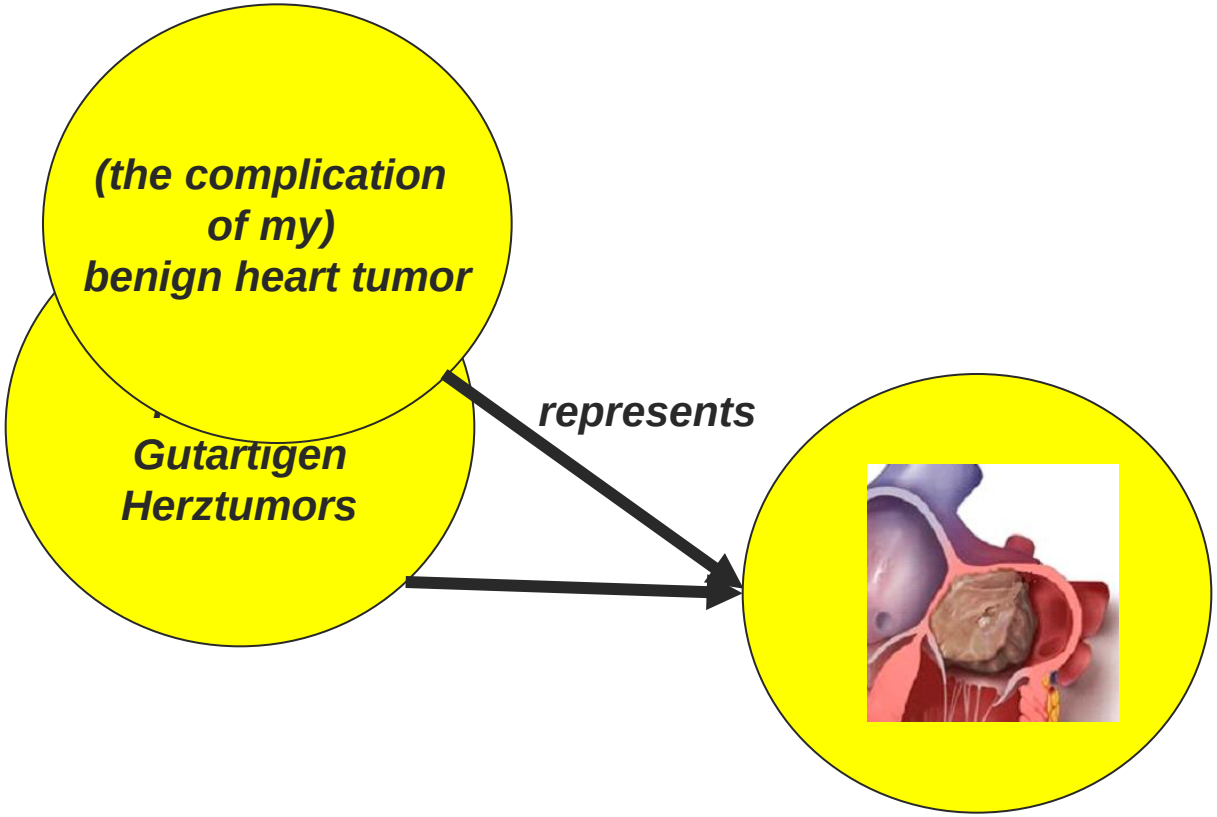
My benign neoplasm of heart



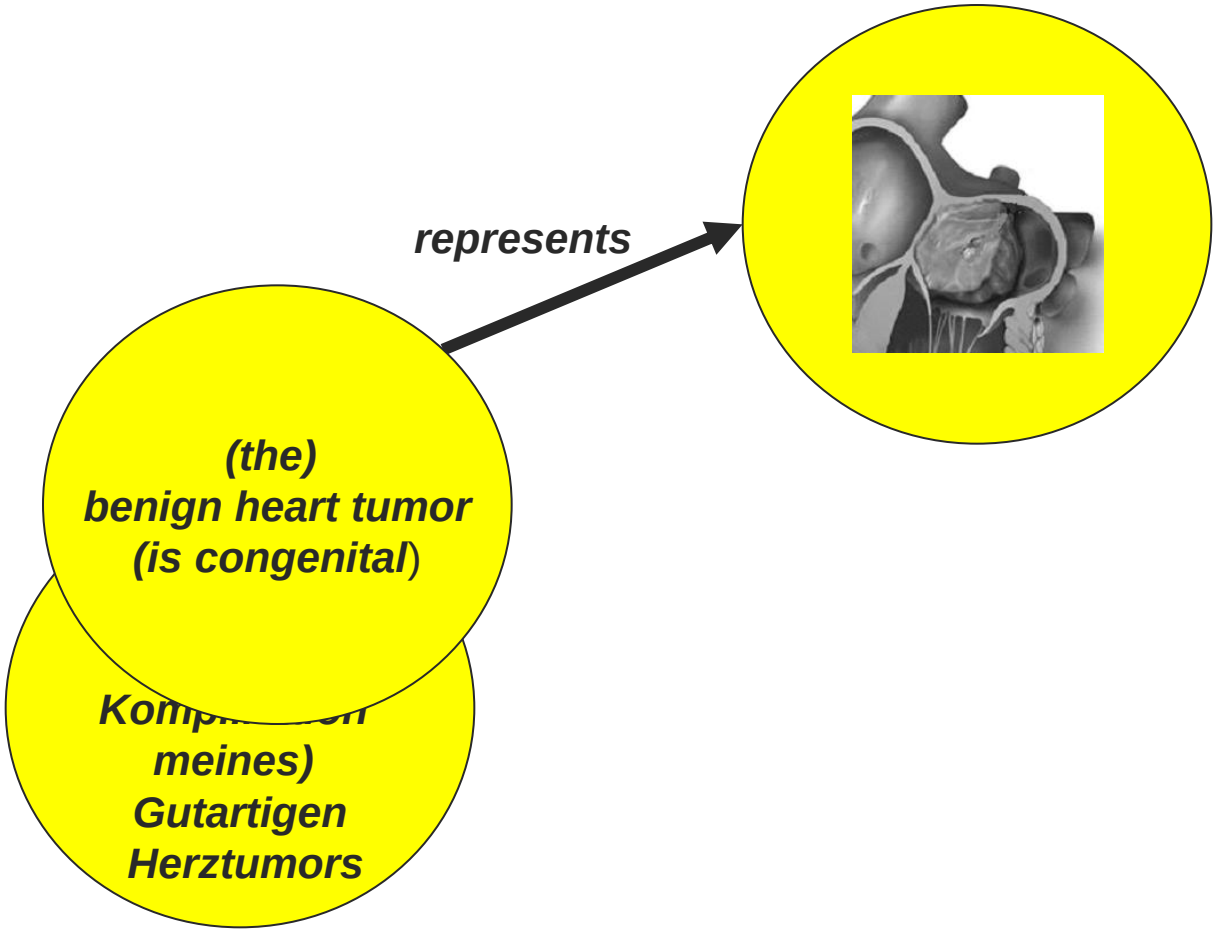
Organizing Entities



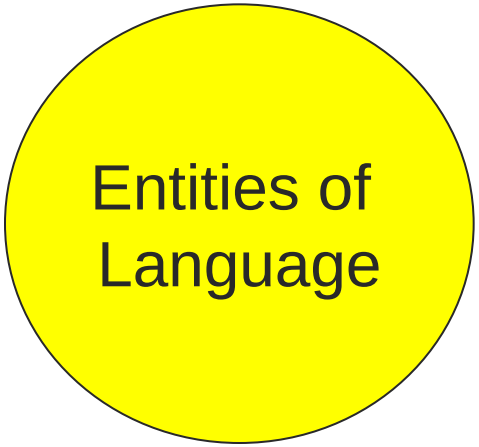
Organizing Entities



Organizing Entities



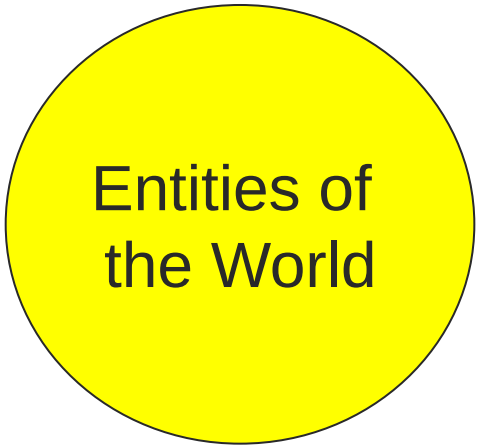
Organizing Entities



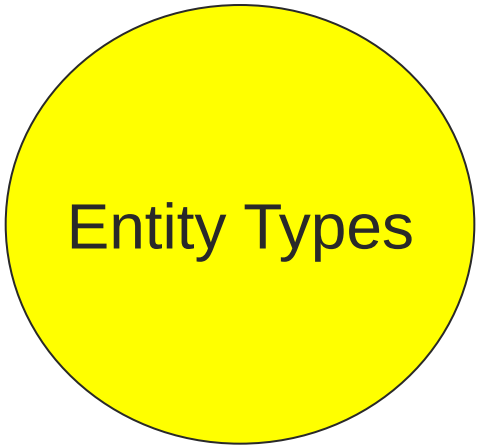
…are stored in dictionaries

Organizing Entities

Database systems store references to…



Organizing Entities



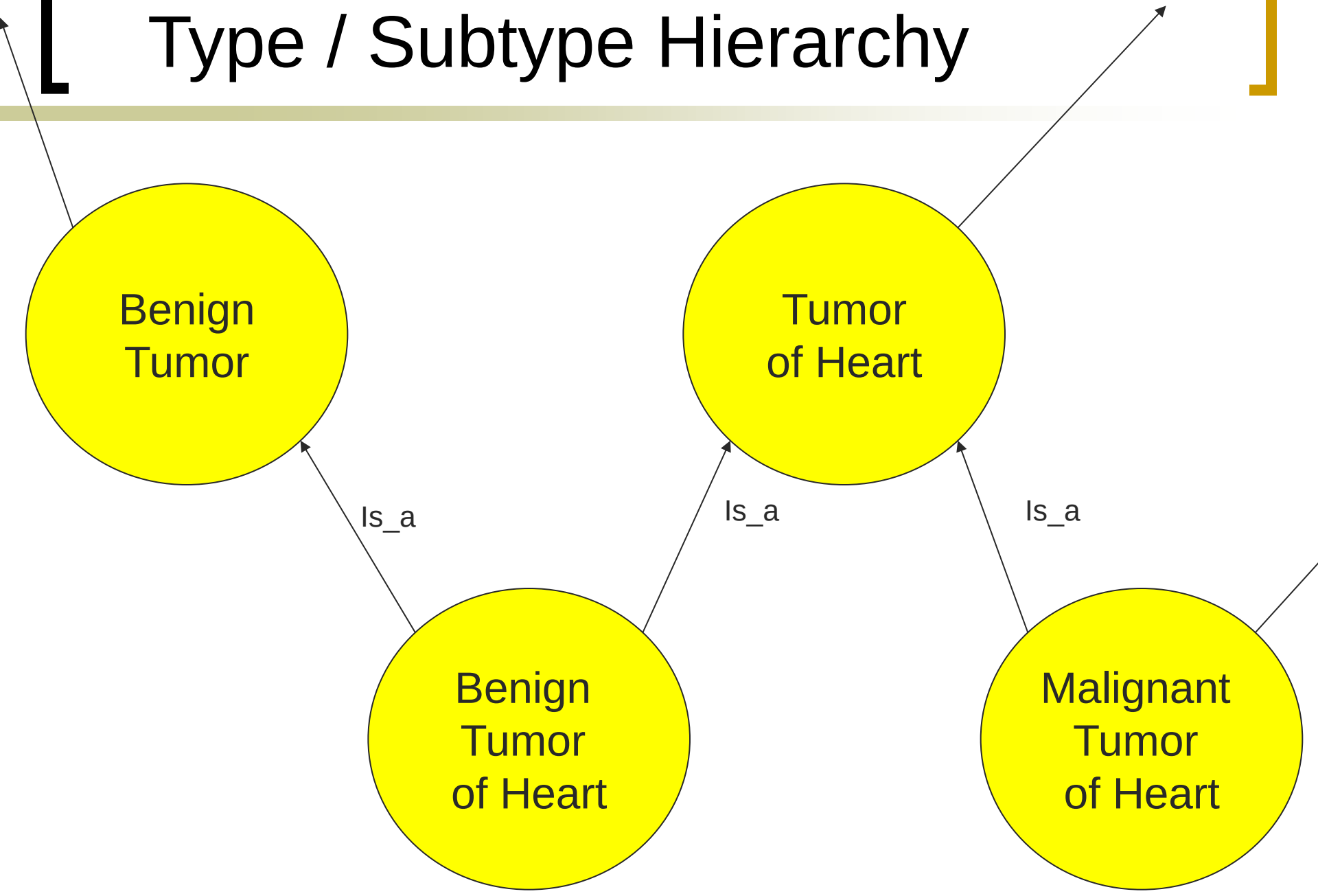
… are organized in ontologies

Hierarchical framework for entity types

- Taxonomy: relates types and subtypes:
  - *is_a (Tumor of Heart, Tumor)* equivalent to:
  - $\forall x$: *instance_of(x, Tumor of Heart)* $\Rightarrow$ *instance_of(x, Tumor)*
- Relations*:
  - *instance_of* relates instances with types, all others relate instances (e.g. *part_of*) or are derived from them (e.g. *is_a*)
- Definitions: formally describe what is always true for all instances of a type
  - $\forall x$: *instance_of(x, Tumor of Heart)* $\Rightarrow$ $\exists y$: *instance_of(y, Heart)* $\wedge$ *has_location(x,y)*

* Smith B, Ceusters W, Kohler J, Kumar A, Lomax J, Mungall CJ, Neuhaus F, Rector A, Rosse C (2005) Relations in Biomedical Ontologies. Genome Biology, 6(5)

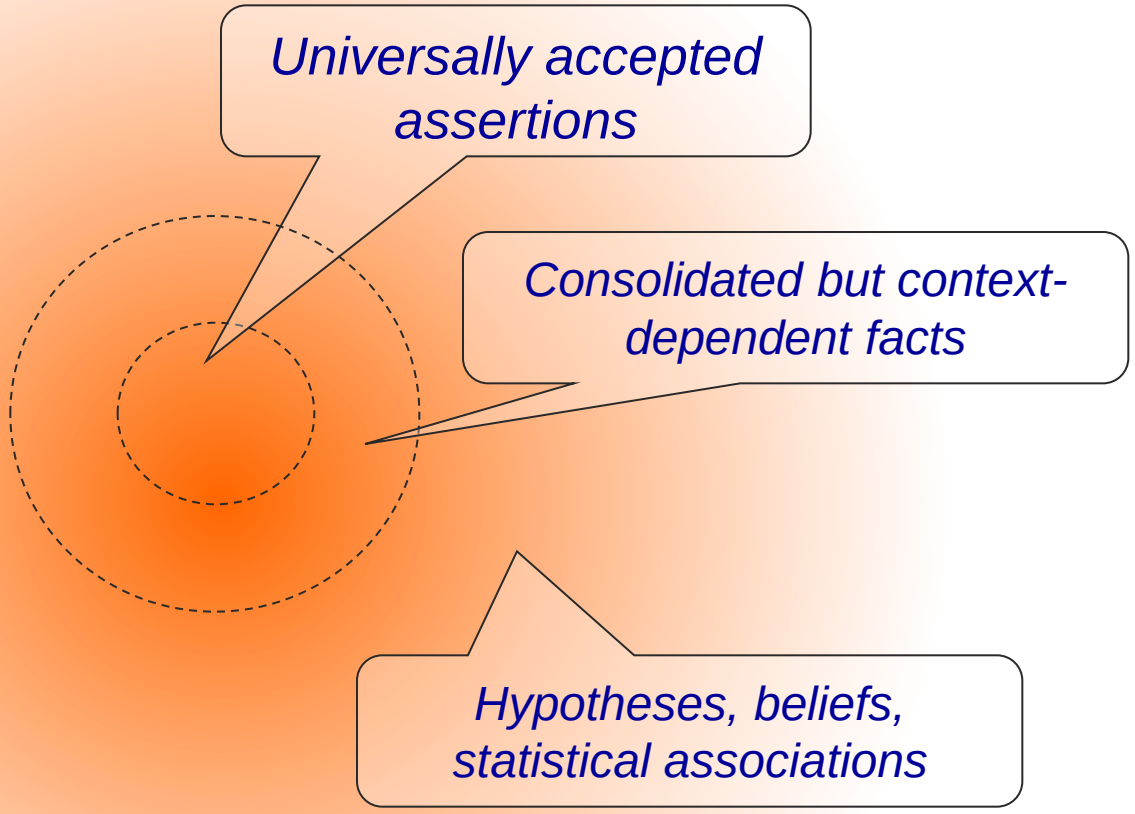
Type / Subtype Hierarchy



III. Ontologies represent “universal truths*”

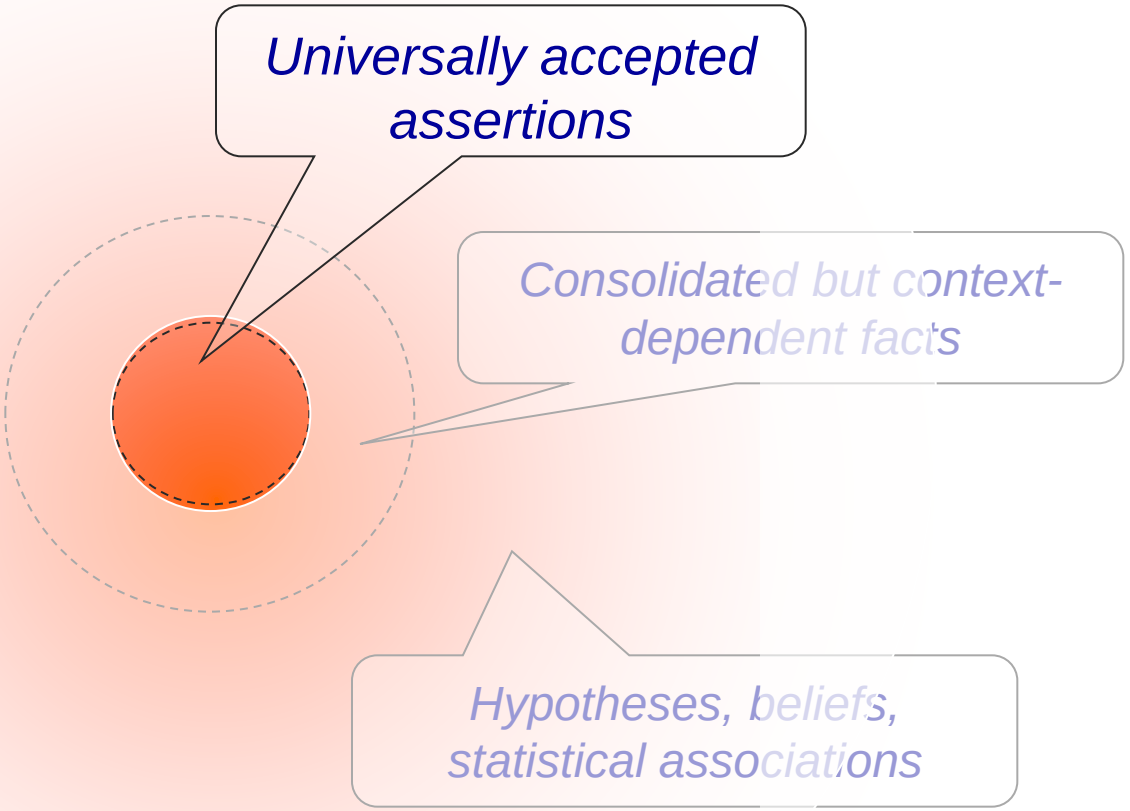
*Assertions that are uncontroversially accepted and very rarely subject to change in a domain, regardless of the philosophical question whether universal truths exist or whether we can know them

Continuum of knowledge



Domain Knowledge

Ontology !



Domain Knowledge

Ontologies: Limitations

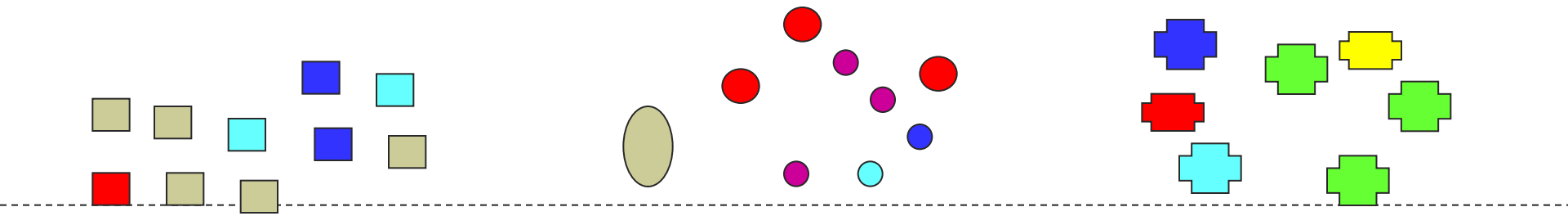
- Represent only shared, uncontroversial meaning of a domain vocabulary
- Makes universal (not probabilistic) statements about instances of a type:
  - All Xs are Ys

    $\forall x: instance\_of(x, X) \Rightarrow instance\_of(x, Y)$
  - For all Xs there is some Y

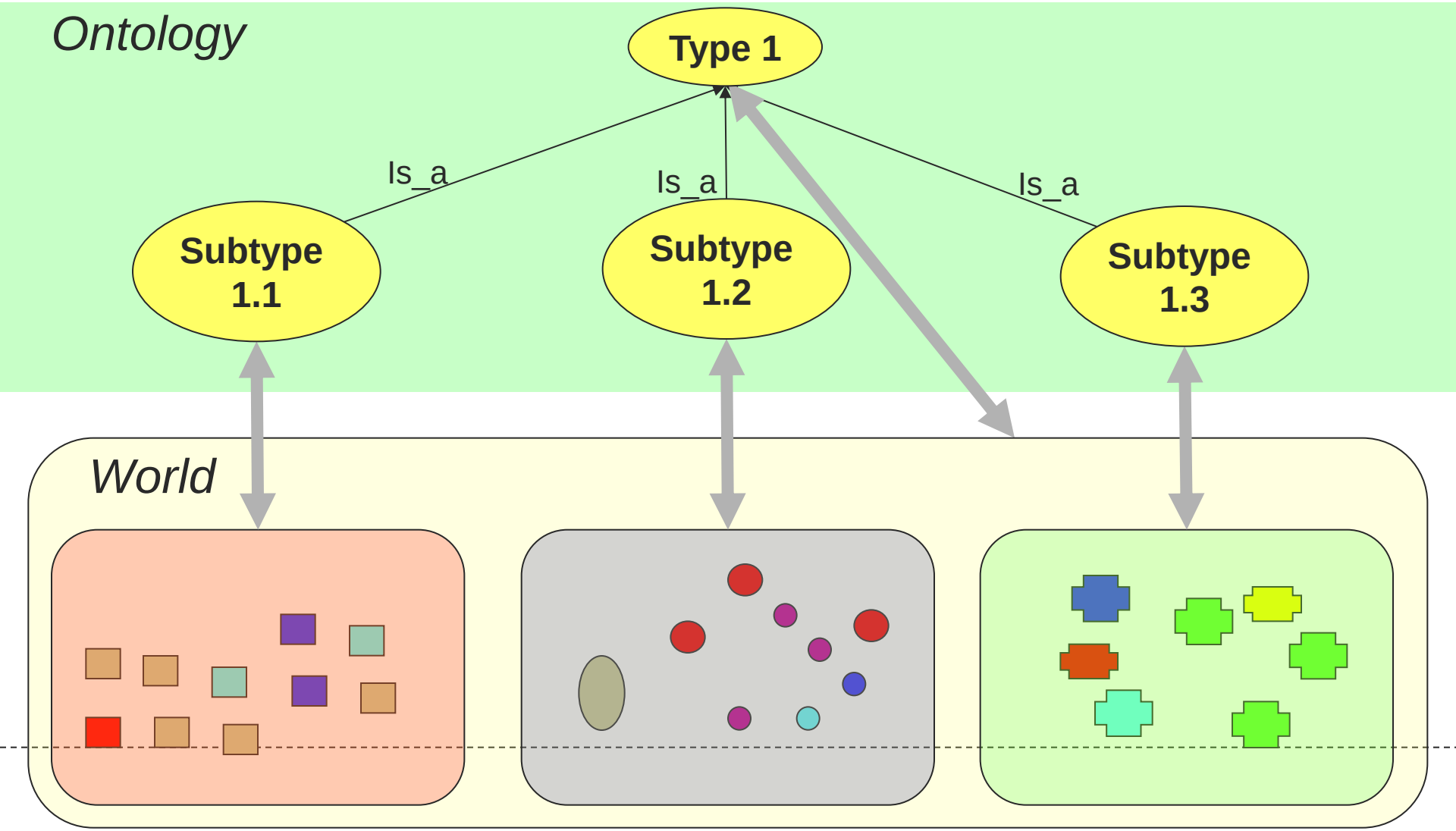
    $\forall x: instance\_of(x, X) \Rightarrow \exists y: \ldots$
- Properties of types are properties of all entities that instantiate these types (strict inheritance)

IV. Ontology types extend to classes of world entities

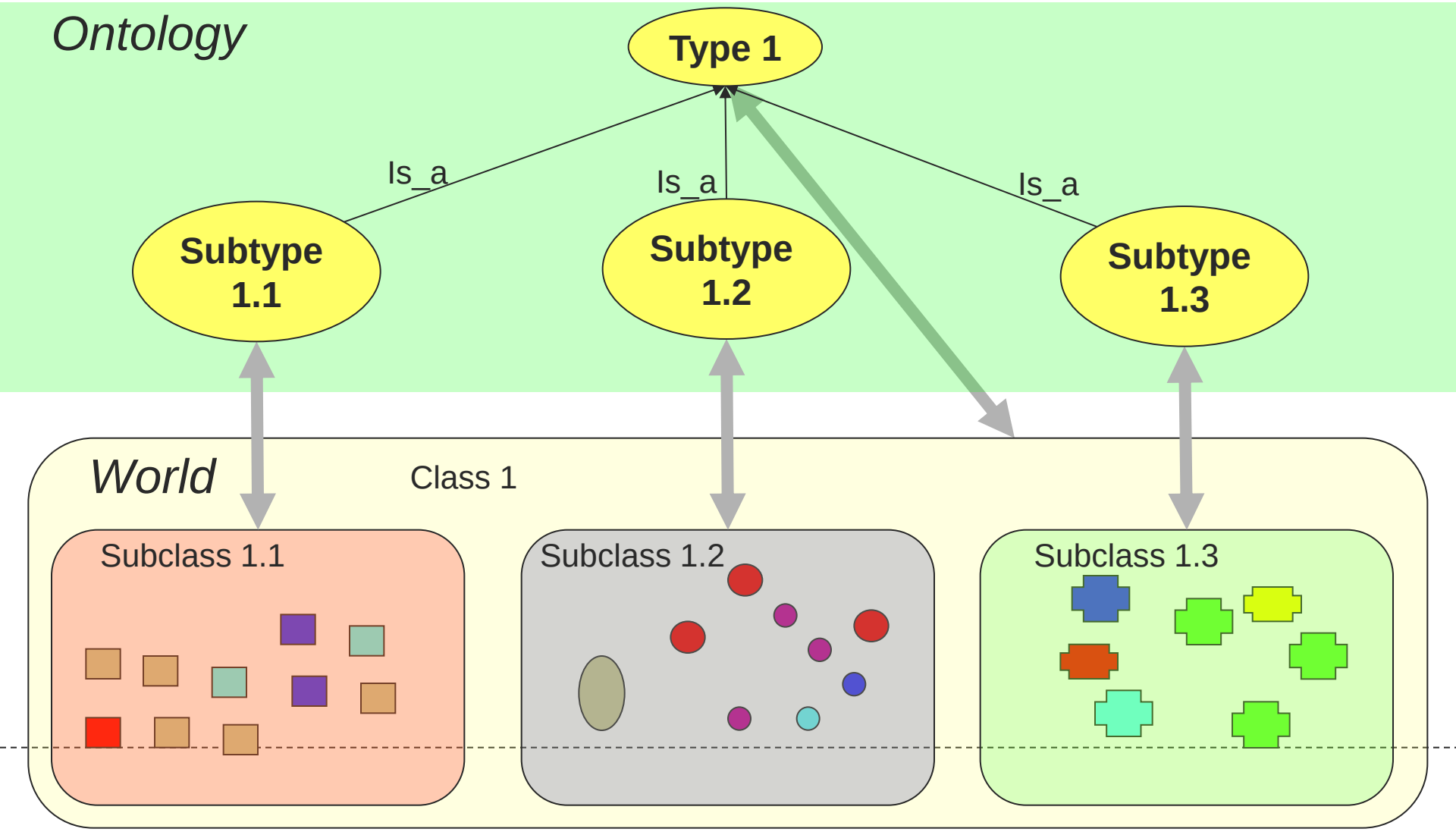
Hierarchies, Types and Classes



Hierarchies, Types and Classes

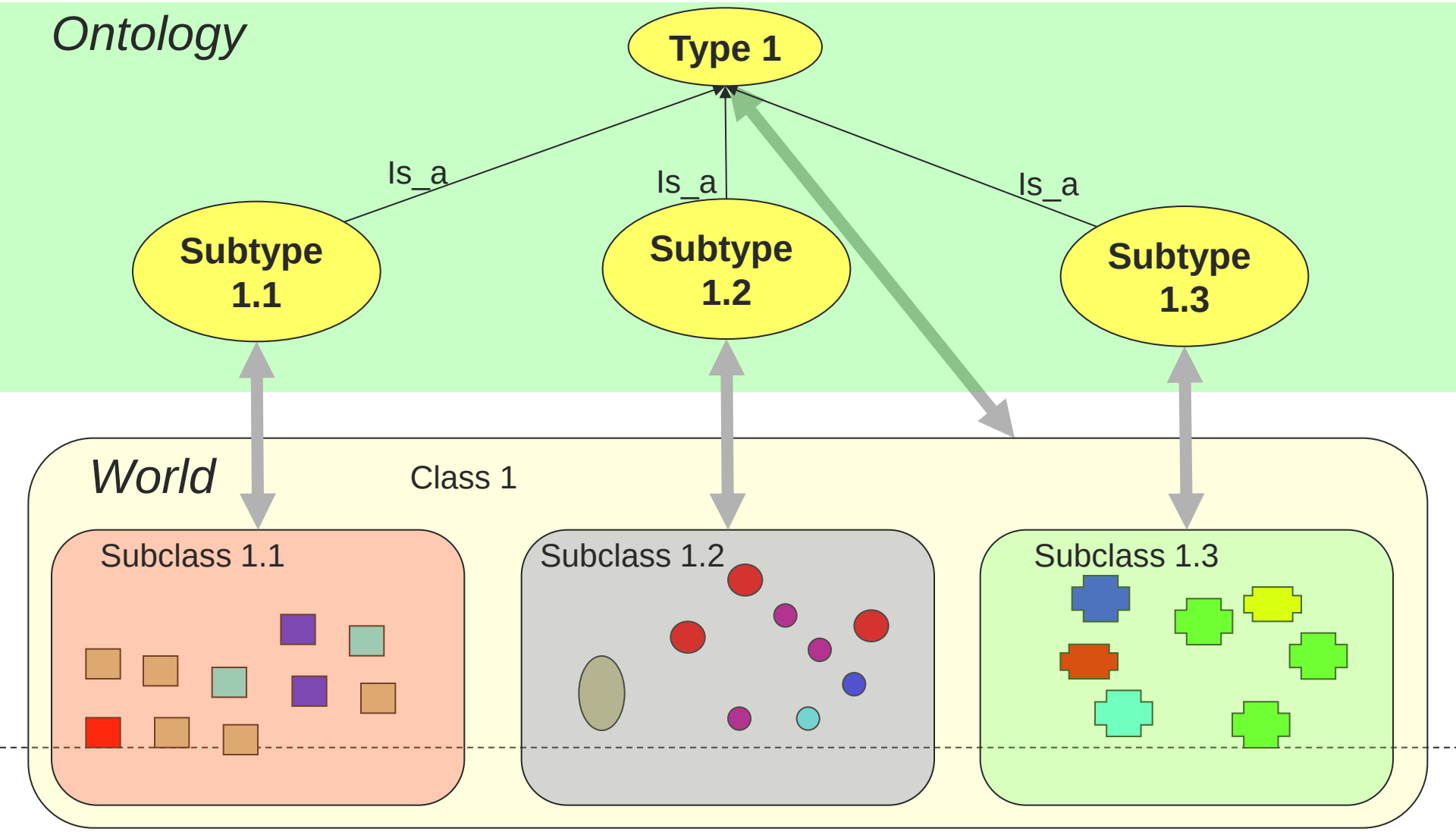


Hierarchies, Types and Classes

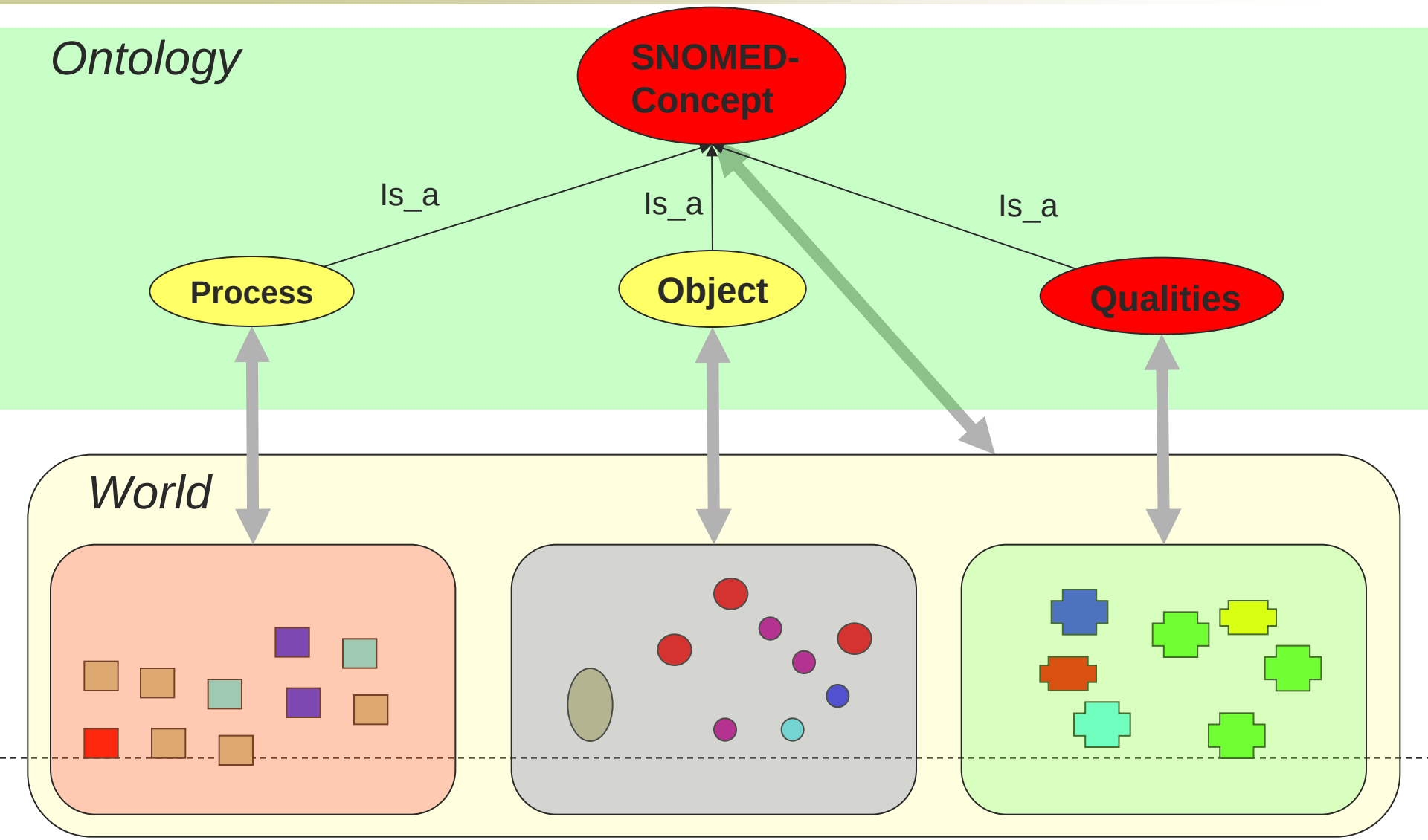


V. Ontologies organize individual entities, not concepts

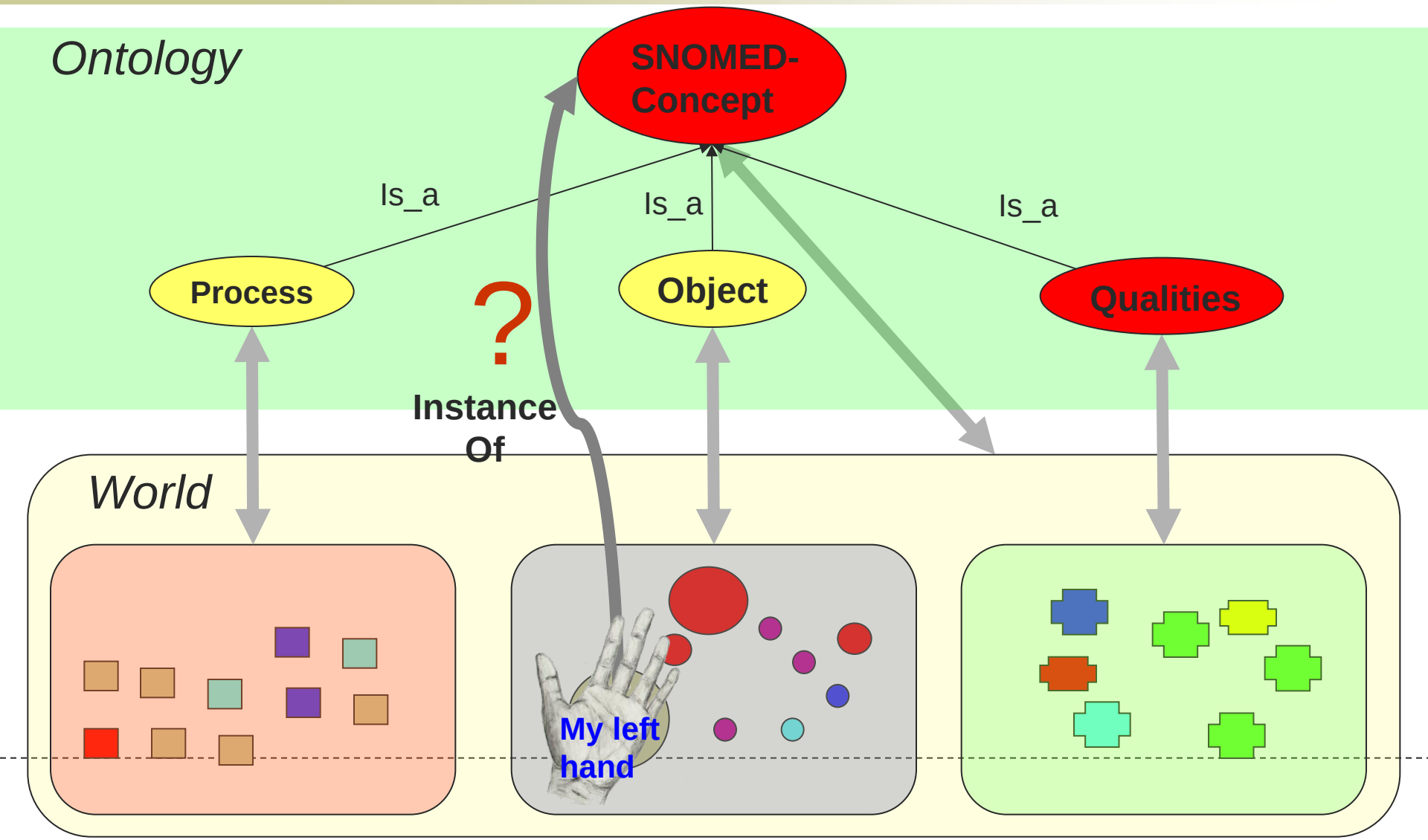
… no direct reference to entities of language or thought



Misconceptions in Ontologies

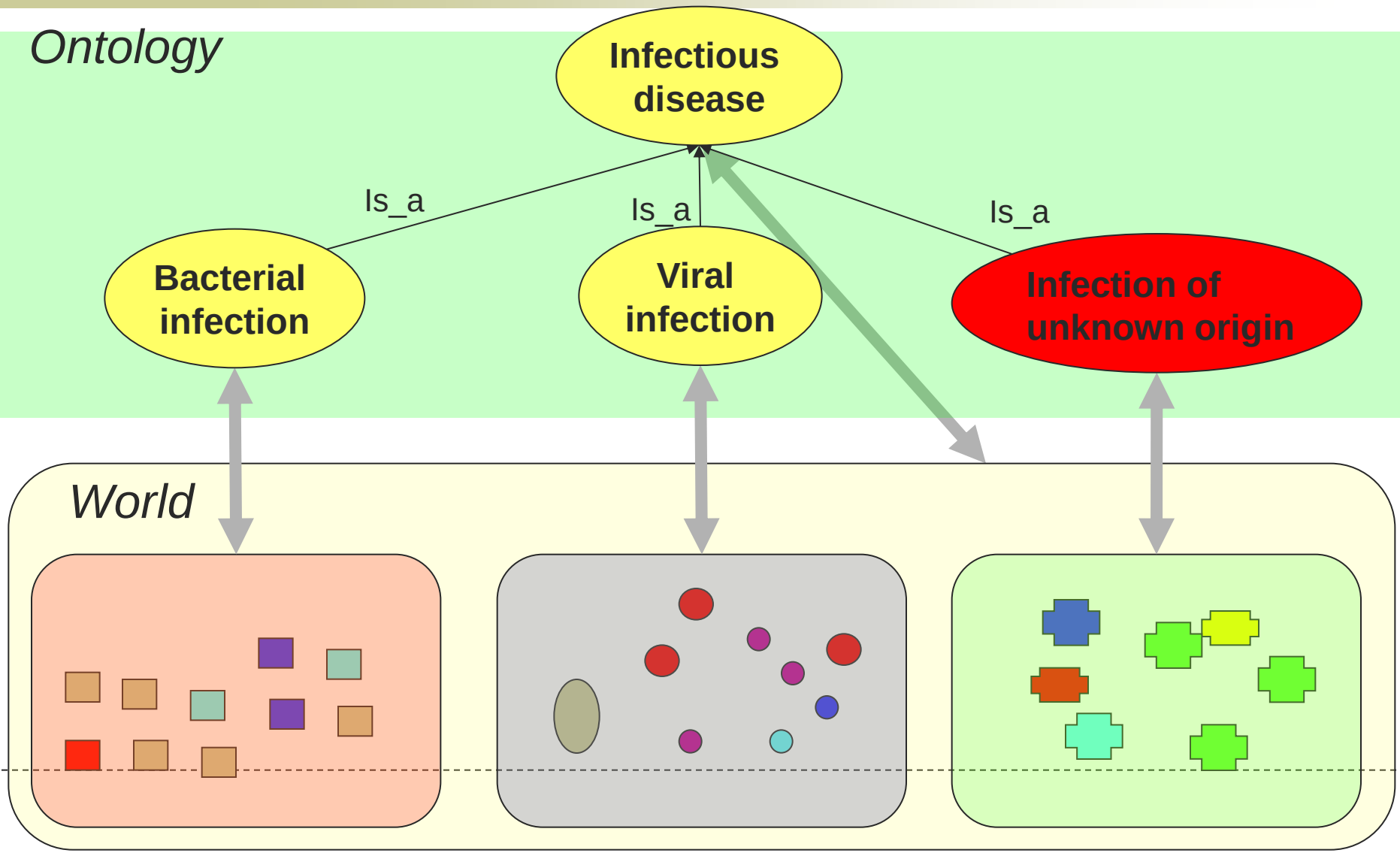


Misconceptions in Ontologies



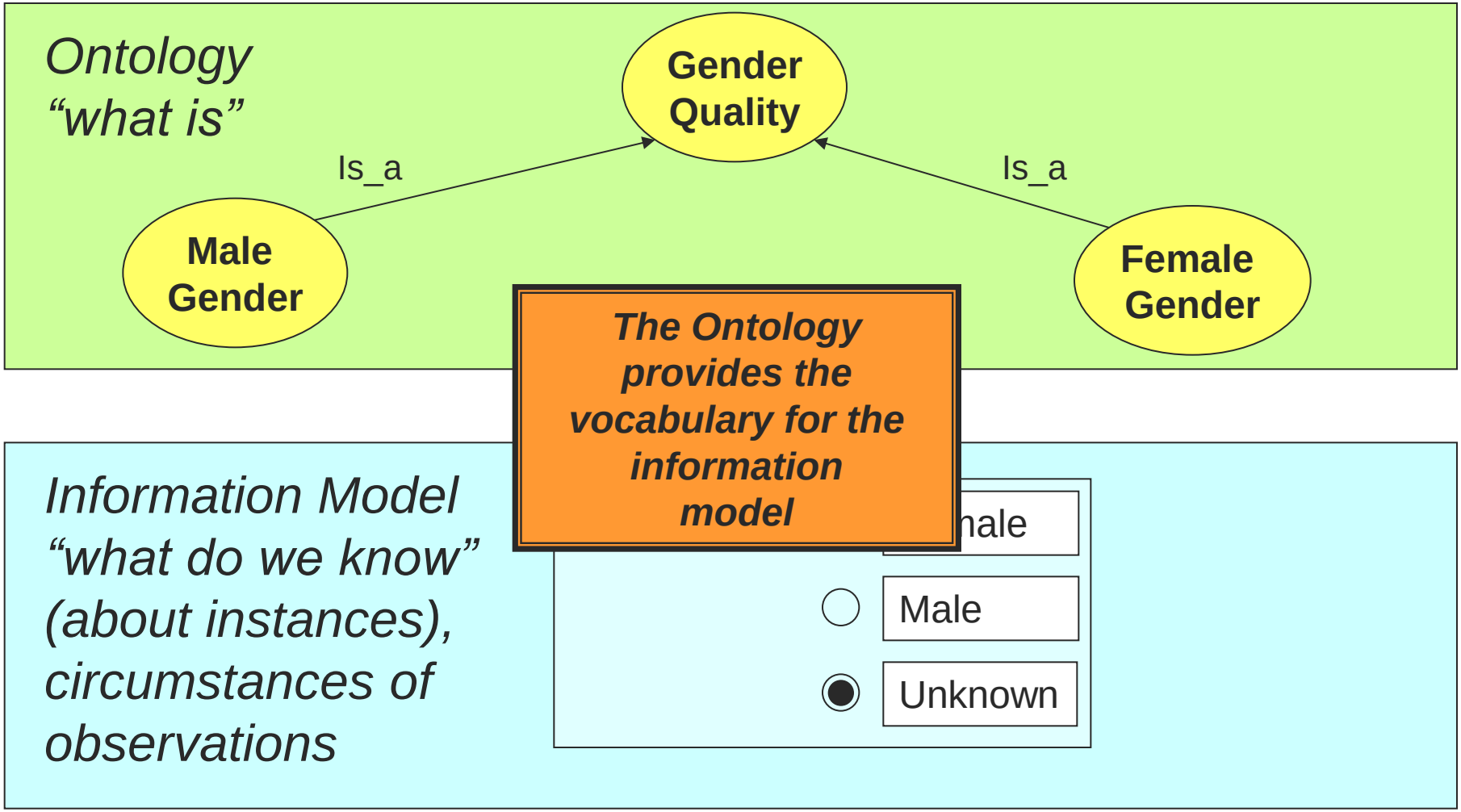
Don’t mix Ontology with Epistemology

O: Bodenreider, B: Smith, and A: Burgun. The ontology-epistemology divide: A case study in medical terminology. In Achille C. Varzi and Laure Vieu, editors, Formal Ontology in Information Systems. Proceedings of the 3rd International Conference - FOIS 2004, pages 185–195. Amsterdam etc.: IOS Press, 2004 .



VI. Ontologies represent what is, information models represent what we know about

Difference
Ontologies / Information Models



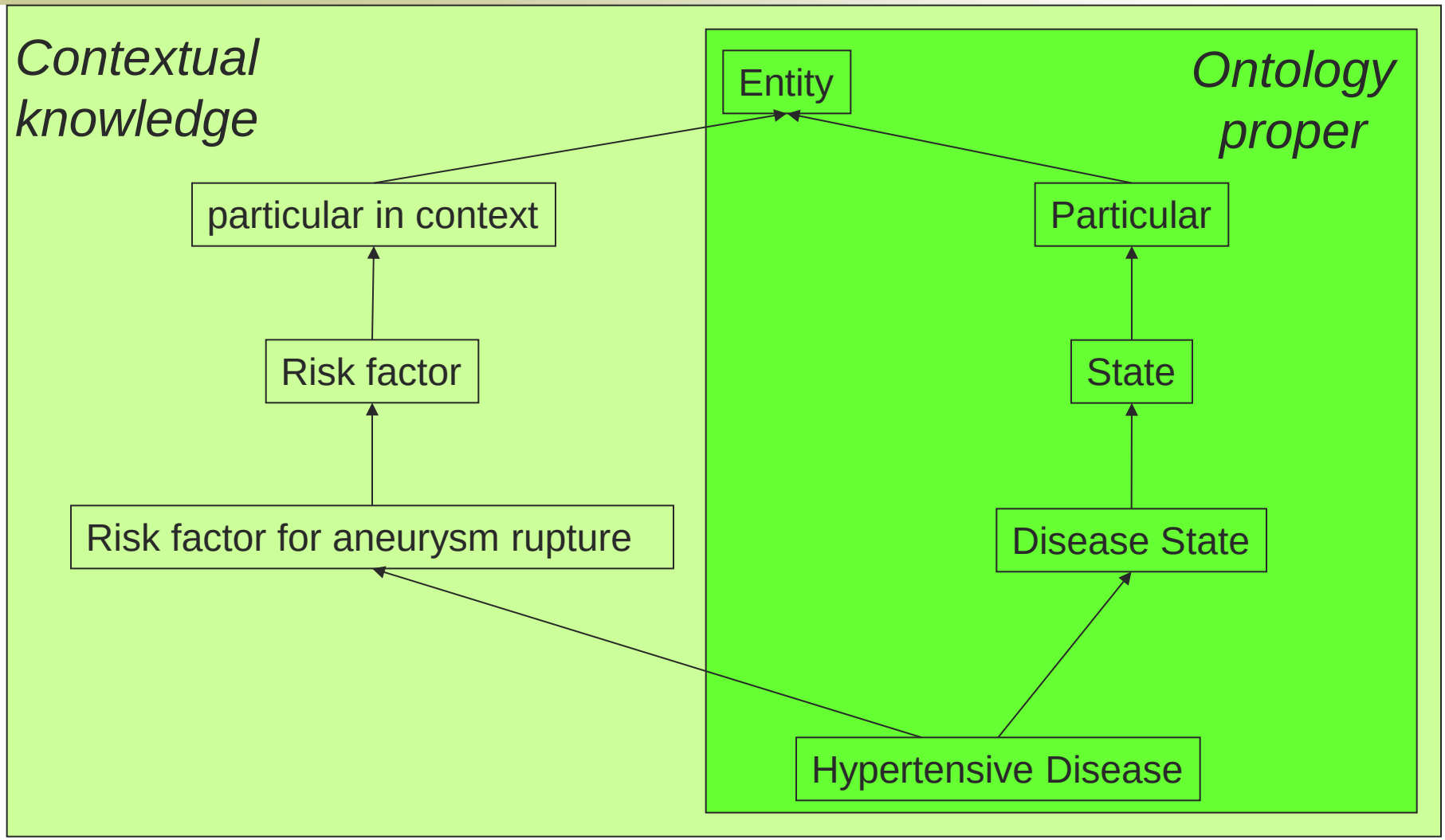
A. Rector, R. Qamar, T. Marley. Binding Ontologies & Coding Systems to Electronic Health Records and Messages. In: Bodenreider O, editor. Formal Biomedical Knowledge Representation (KR-MED 2006) CEUR; 2006. p. 11-19.

VII. Practical requirements may justify controlled deviations from the “true path”

Ontology in practice: compromises required

- Example: Risk factors:
  - A is a risk factor for B: Being risk factor for B is a role A plays under certain circumstances in certain populations.
  - It means that there is a statistical dependency of the incidence of B from on the presence of A in a population
  - It is not an ontological property of A: not every instance of A is associated to some B

@neurist: Ontology in practice: Risk factor workaround



S. Hanser, M. Boeker, K. Kumpf, P. Bijlenga, S. Schulz. Design of an Ontology on Cerebral Aneurysms: Representing the Conceptual Space of the @neurIST Project. Medinfo 2007 Congress, 20-24 August 2007, Brisbane, Australia. Accepted for publication.

Beyond the expressiveness of Formal Ontologies

- Probabilistic knowledge:
  - Presence of jaundice in 95% cases of hepatitis
- Default assertions:
  - Adult humans have 32 teeth
- Population-based assertions
  - Hayfever is a common seasonal disorder in the NL
- Dispositions
  - Gleevec® cures CML
  - Aspirin® irritates the Gastric Mucosa

Beyond the expressiveness of ontology languages for formal reasoning (e.g. OWL-DL)

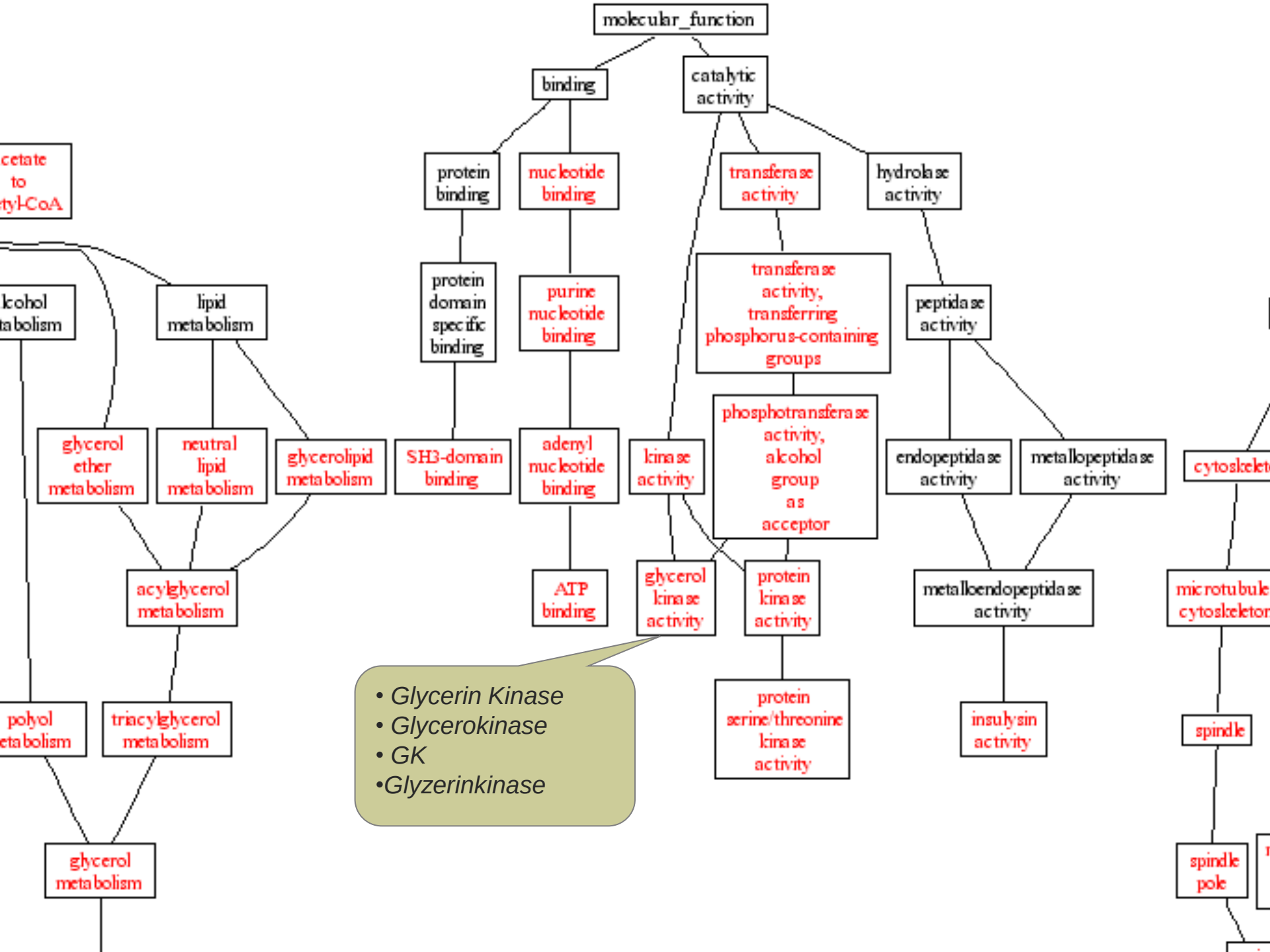
Logical Foundations of Ontologies

DL	FOL
$a : C$	$C(a)$
$\langle a, b \rangle : P$	$P(a, b)$
$C \sqsubseteq D$	$\forall x.C(x) \rightarrow D(x)$
$P^+ \sqsubseteq P$	$\forall x, y, z.(P(x, y) \wedge P(y, z)) \rightarrow P(x, z)$
$\top \sqsubseteq \leqslant 1\, P$	$\forall x, y, z.(P(x, y) \wedge P(x, z)) \rightarrow y = z$
$P \equiv Q^-$	$\forall x, y.P(x, y) \iff Q(y, x)$
$C_1 \sqcap \ldots \sqcap C_n$	$C_1(x) \wedge \ldots \wedge C_n(x)$
$C_1 \sqcup \ldots \sqcap C_n$	$C_1(x) \vee \ldots \vee C_n(x)$
$\neg C$	$\neg C(x)$
$\{a_1, \ldots, a_n\}$	$x = a_1 \vee \ldots \vee x = a_n$
$\exists P.C$	$\exists y.(P(x, y) \wedge C(y))$
$\forall P.C$	$\forall y..(P(x, y) \rightarrow C(y))$
$\geqslant n\, P.C$	$\exists y_1, \ldots, y_n.\, \bigwedge_{1 \leqslant i \leqslant n}(P(x, y_i) \wedge C(y_i)) \wedge \bigwedge_{1 \leqslant i < n, i < j \leqslant n} y_i \neq y_j$
$\leqslant (n-1)\, P.C$	$\forall y_1, \ldots, y_n.(\bigwedge_{1 \leqslant i \leqslant n}(P(x, y_i) \wedge C(y_i))) \rightarrow (\bigvee_{1 \leqslant i < n, i < j \leqslant n} y_i = y_j)$

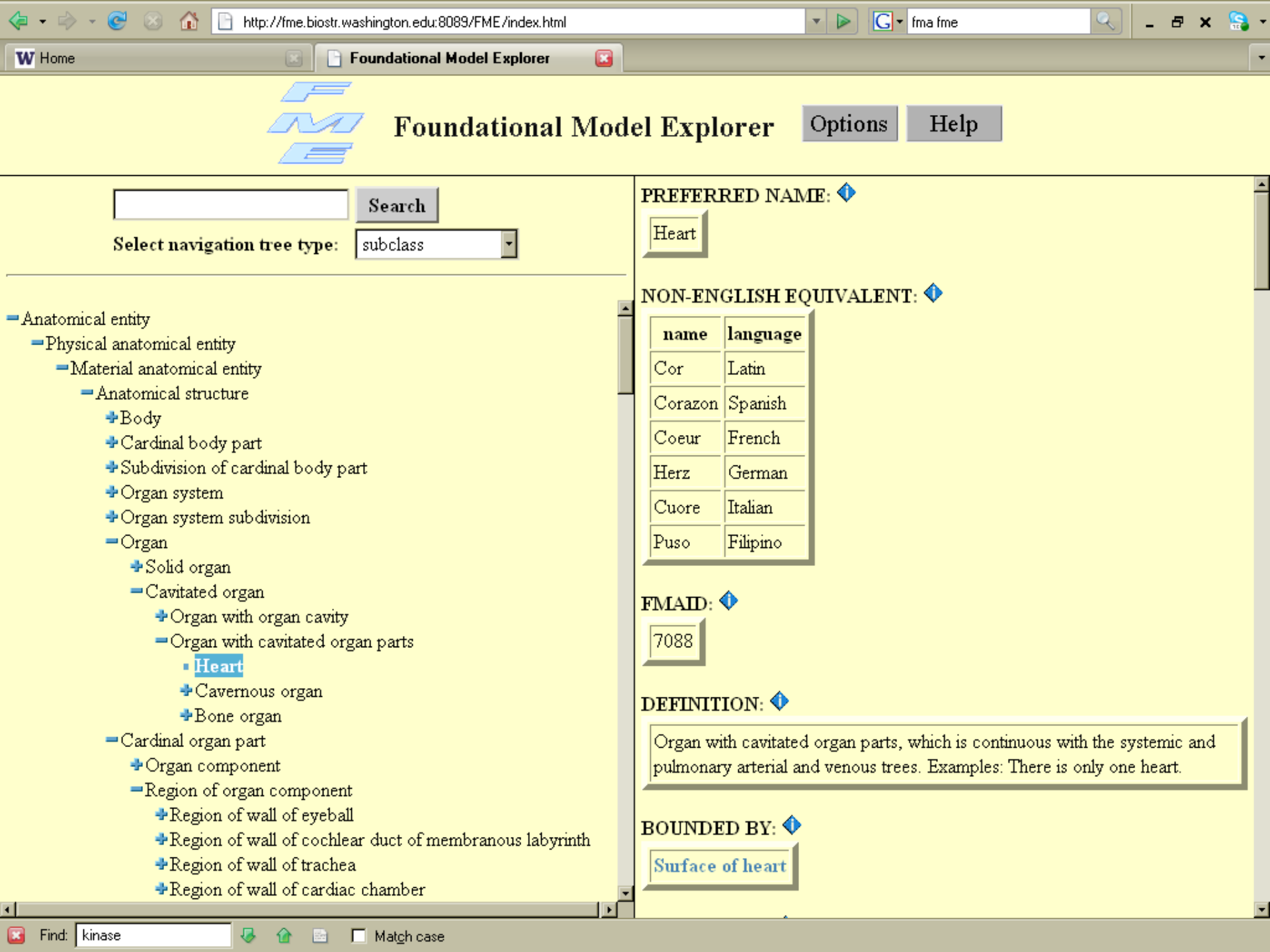
VIII. Ontologies need to be linked to vocabularies

Ontologies and Vocabularies

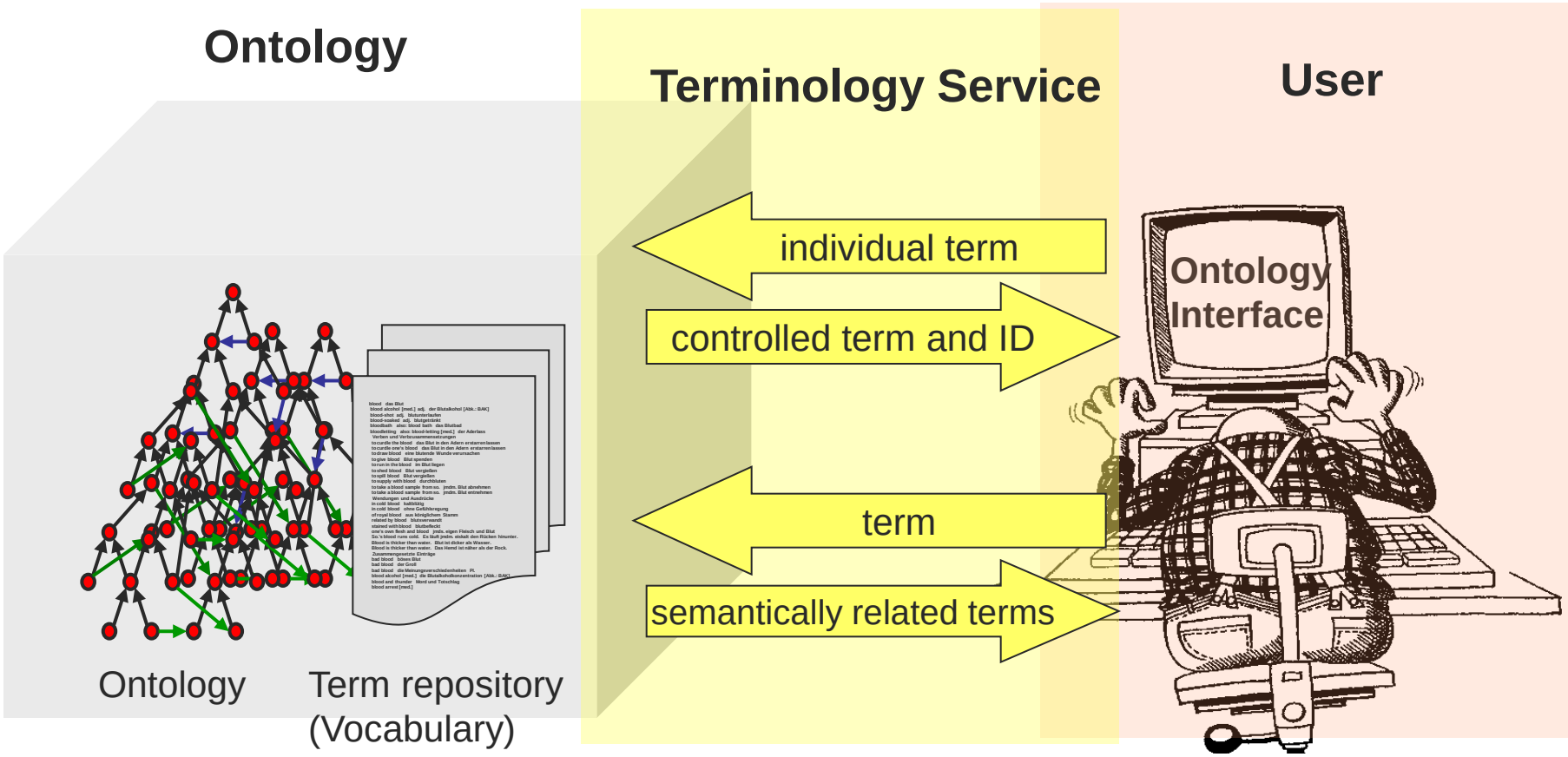
- Vocabularies: Lists of terms with defined meaning in a domain
- Human-readable Labels are commonly used for unambiguously characterizing classes in an ontology
- Labels should not be mistaken for domain terms:
  - Domain terms may be ambiguous
  - Synonyms are common


molecular_function
binding
catalytic activity
protein binding
nucleotide binding
transferase activity
hydrolase activity
protein domain specific binding
purine nucleotide binding
transferase activity, transferring phosphorus-containing groups
peptidase activity
SH3-domain binding
adenyl nucleotide binding
kinase activity
phosphotransferase activity, alcohol group as acceptor
endopeptidase activity
metallopeptidase activity
ATP binding
glycerol kinase activity
protein kinase activity
metalloendopeptidase activity
protein serine/threonine kinase activity
insulysin activity
lipid metabolism
glycerol ether metabolism
neutral lipid metabolism
glycerolipid metabolism
acylglycerol metabolism
triacylglycerol metabolism
glycerol metabolism
• Glycerin Kinase
• Glycerokinase
• GK
•Glyzerinkinase
spindle
spindle pole

IX. Ontology users need not to see the whole “engine”


http://fme.biostr.washington.edu:8089/FME/index.html
fma fme
Home
Foundational Model Explorer
Foundational Model Explorer
Options
Help
Search
Select navigation tree type:
subclass
Anatomical entity
Physical anatomical entity
Material anatomical entity
Anatomical structure
Body
Cardinal body part
Subdivision of cardinal body part
Organ system
Organ system subdivision
Organ
Solid organ
Cavitated organ
Organ with organ cavity
Organ with cavitated organ parts
Heart
Cavernous organ
Bone organ
Cardinal organ part
Organ component
Region of organ component
Region of wall of eyeball
Region of wall of cochlear duct of membranous labyrinth
Region of wall of trachea
Region of wall of cardiac chamber
PREFERRED NAME:
Heart
NON-ENGLISH EQUIVALENT:
name
language
Cor
Latin
Corazon
Spanish
Coeur
French
Herz
German
Cuore
Italian
Puso
Filipino
FMAID:
7088
DEFINITION:
Organ with cavitated organ parts, which is continuous with the systemic and pulmonary arterial and venous trees. Examples: There is only one heart.
BOUNDED BY:
Surface of heart
Find:
kinase
Match case

The user perspective: Ontology as logic based terminology service



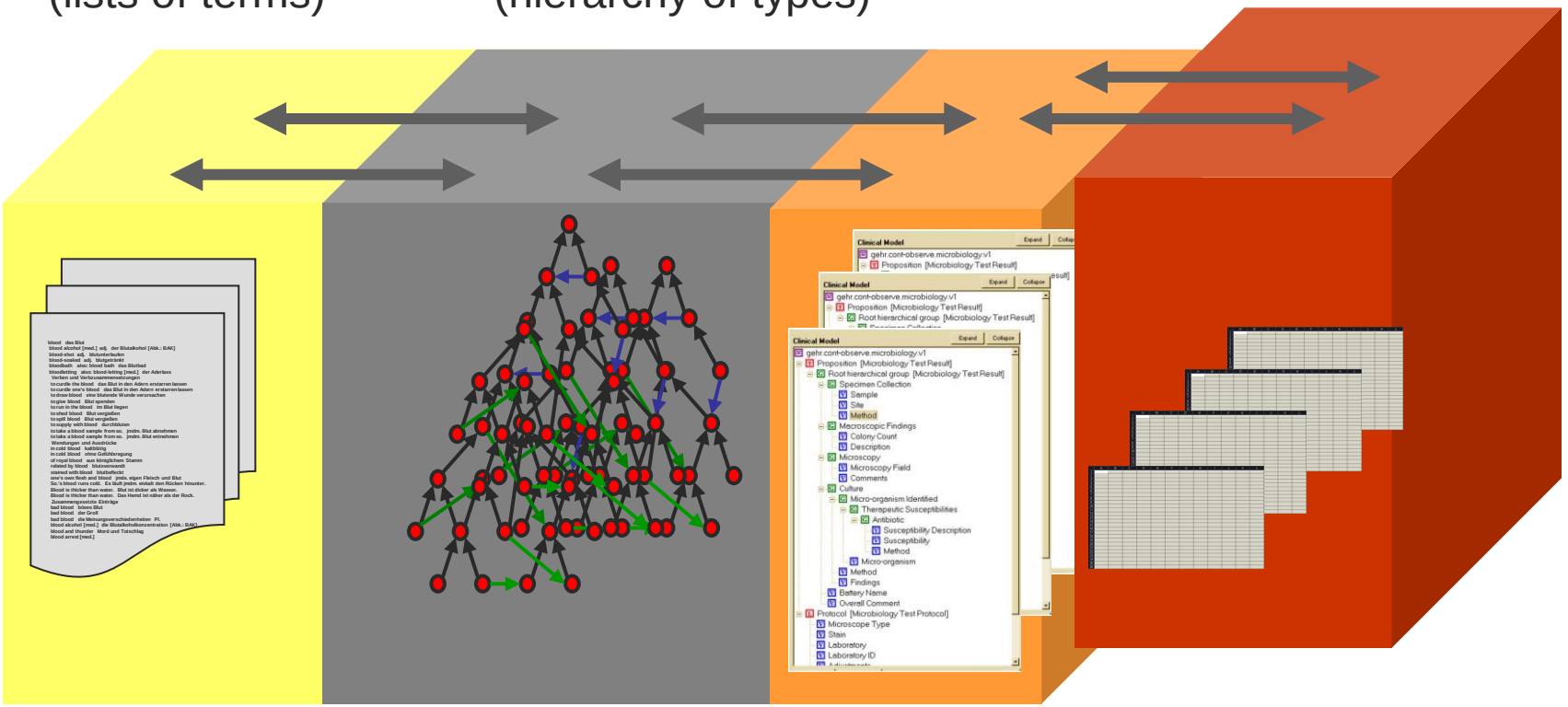
X. Ontology in practice should support tailored logic-based terminology services

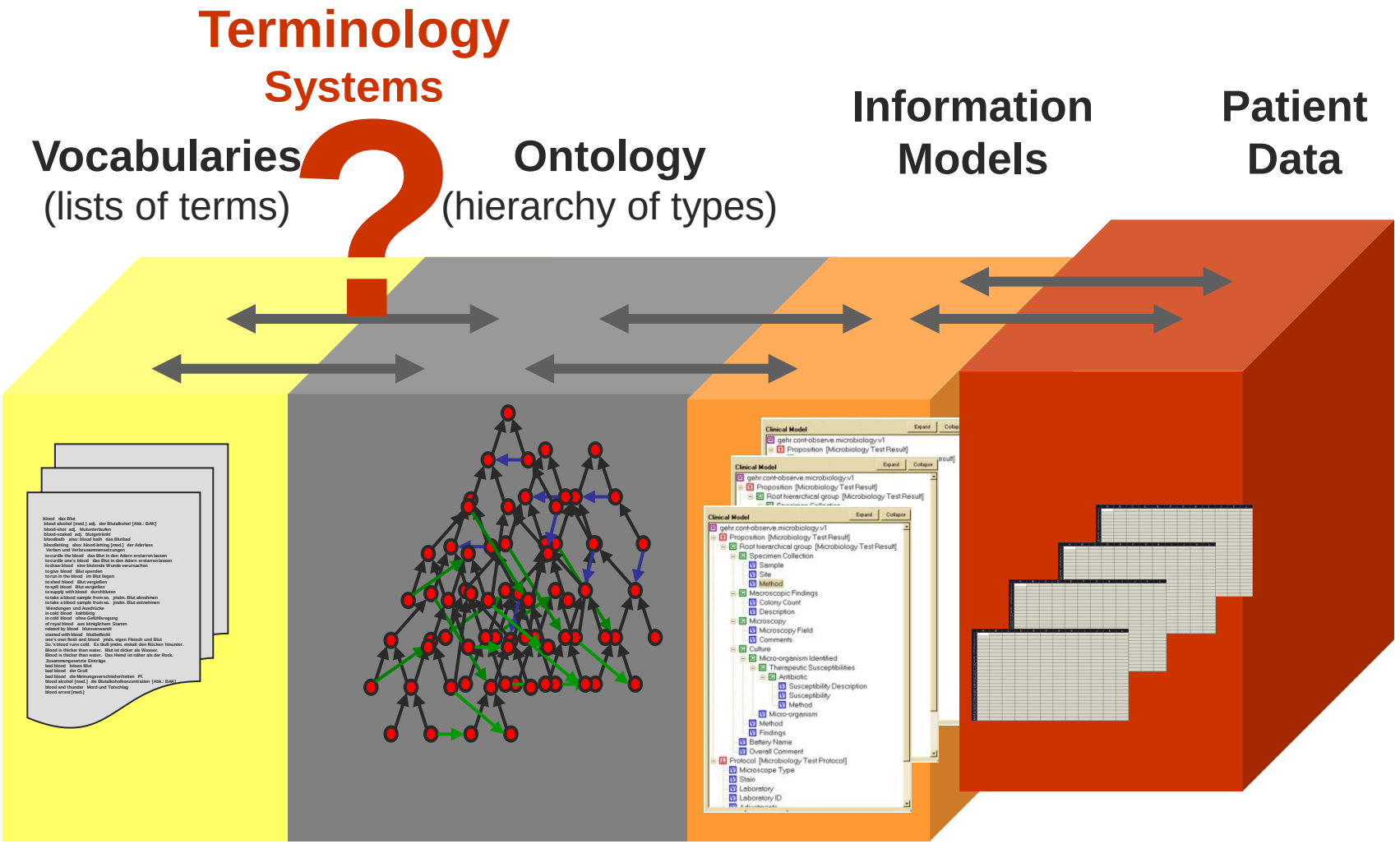
Vocabularies
(lists of terms)

Ontology
(hierarchy of types)

Information Models

Patient Data




Terminology
Systems
Vocabularies
(lists of terms)
Ontology
(hierarchy of types)
Information
Models
Patient
Data
Clinical Model
gehr.cont-observe.microbiology.v1
Proposition [Microbiology Test Result]
Root hierarchical group [Microbiology Test Result]
Specimen Collection
Sample
Site
Method
Microscopic Findings
Colony Count
Description
Microscopy
Microscopy Field
Comments
Culture
Micro-organism Identified
Therapeutic Susceptibilities
Antibiotic
Susceptibility Description
Susceptibility
Method
Micro-organism
Method
Findings
Battery Name
Overall Comment
Protocol [Microbiology Test Protocol]
Microscope Type
Stain
Laboratory
Laboratory ID

Open questions: Bringing Ontologies and Terminology Systems Together

- Which use cases require (formal) ontologies
- In which cases informal terminology systems are sufficient?
- Which cases require both ?
- Can existing terminologies be ontologized?
- Can terminologies and ontologies co-exist ?

Ten principles

I. **Terminology systems provide semantic reference**
II. **Ontologies are hierarchies of semantic types that support the organization of domain entities**
III. **Ontologies represent universal truths**
IV. **Ontology types extend to classes of world entities**
V. **Ontologies organize individual entities, not concepts**
VI. **Ontologies represent what is, information models represent what we know about**
VII. **Practical requirements may justify controlled deviations from the “true path”**
VIII. **Ontologies need to be linked to vocabularies**
IX. **Ontology users need not to see the whole “engine”**
X. **Ontology should provide tailored terminology services**

Thank you for your attention !

stschulz@uni-freiburg.de