Universitätsspital Basel

Clinical research seminar SNOMED CT
University Hospital Basel, May 24th, 2018

Stefan Schulz
Medical University of Graz (Austria)

purl.org/steschu

Annotating clinical narratives with SNOMED CT

The thorny way towards interoperability of clinical routine data

Focus of the talk

- What is semantic interoperability
- Analysis of manual SNOMED CT text annotations
- Inter-annotator agreement (reliability)
- Reasons for inter-annotator disagreement
- Discussion: how to improve agreement
- Conclusions / Outlook

- Assumption:
  better data reliability
  -> better semantic interoperability of clinical data

Focus of the talk

- **What is semantic interoperability**
- Analysis of manual SNOMED CT text annotations
- Inter-annotator agreement (reliability)
- Reasons for inter-annotator disagreement
- Discussion: how to improve agreement
- Conclusions / Outlook

- Assumption:
  better data reliability
  -> better semantic interoperability of clinical data

What is semantic interoperability?

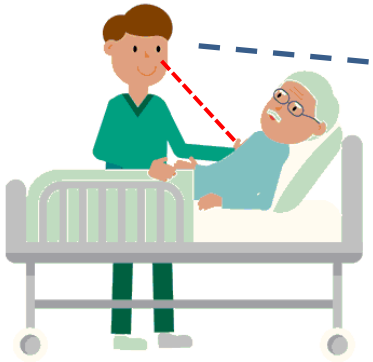
What is semantic interoperability?

Observation

Documentation

Information System A

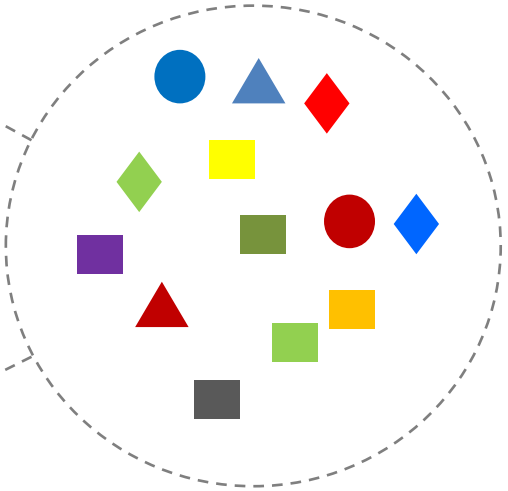
D_A



Documentation

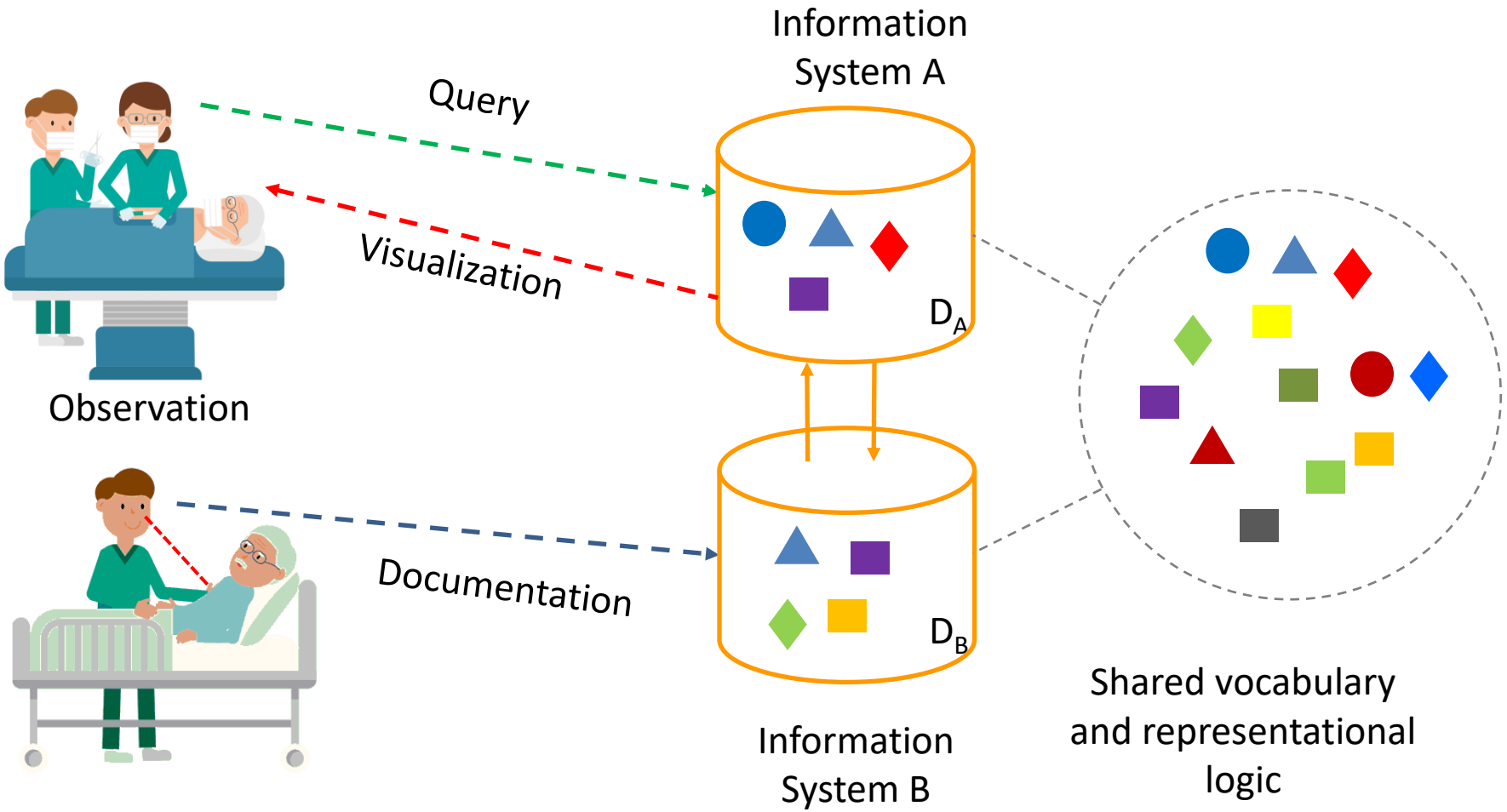
D_B

Information System B

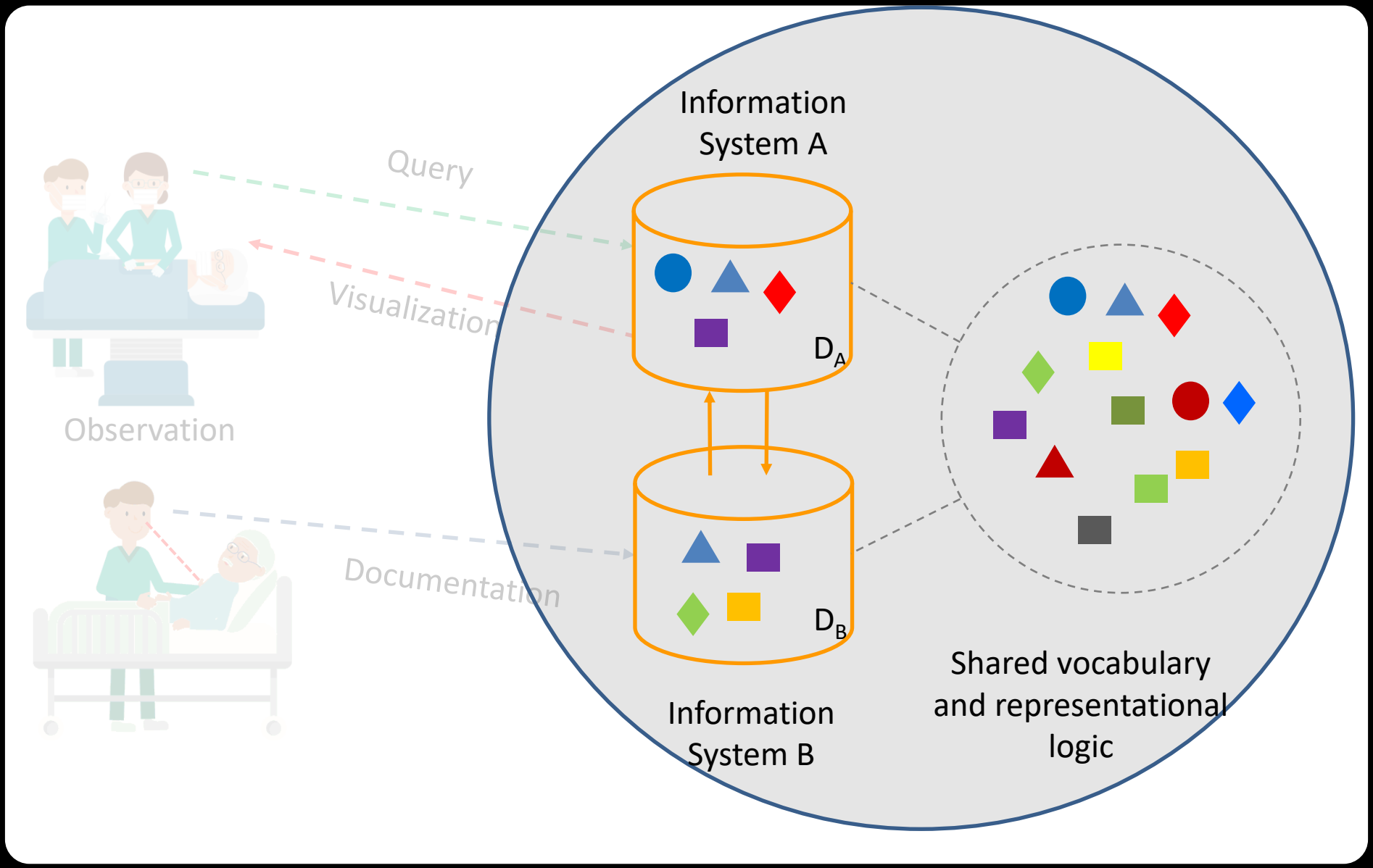


Shared vocabulary and representational logic

What is semantic interoperability?

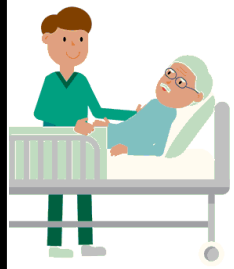


How to validate Interoperability Infrastructure ?

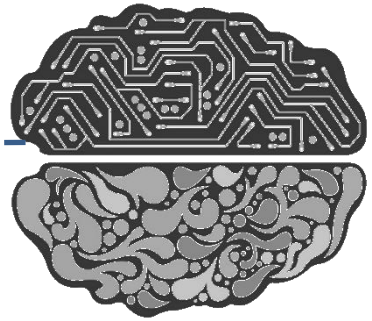


Annotating clinical narratives with SNOMED CT

Coding



observation



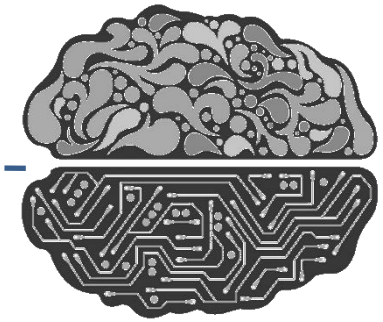
standardised representation
SNOMED CT

Annotation

observation, interpretation

Resektat nach Whipple": Ein noch nicht eröffnetes Resektat, bestehend aus einem distalen Magen mit einer kleinen Kurvaturlänge von 9,5 cm und einer großen Kurvaturlänge von 13,5 cm, sowei einem duodenalen Anteil von 14 cm Länge. 2 cm aboral des Pylorus zeigt die Dünndarmwandung eine sanduhrartige ...

human language representation



standardised representation
SNOMED CT

Focus of the talk

- What is semantic interoperability
- **Analysis of manual SNOMED CT text annotations**
- Inter-annotator agreement (reliability)
- Reasons for inter-annotator disagreement
- Discussion: how to improve agreement
- Conclusions / Outlook

- Assumption:
  better data reliability
  -> better semantic interoperability of clinical data

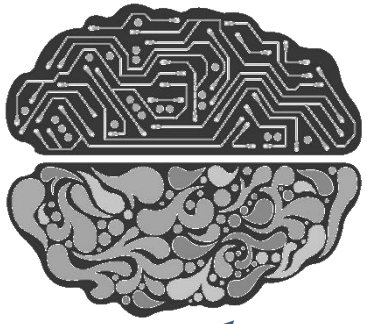
Annotations for SNOMED CT validation

Annotation experiment to validate SNOMED CT

Starting point:
Clinical document corpus

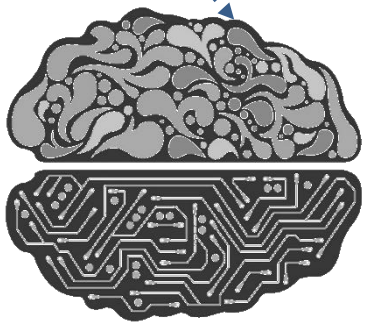
Resektat nach Whipple[n]: Ein noch nicht eröffnetes Resektat, bestehend aus einem distalen Magen mit einer kleinen Kurvaturlänge von 9,5 cm und einer großen Kurvaturlänge von 13,5 cm, sowei einem duodenalen Anteil von 14 cm Länge. 2 cm aboral des Pylorus zeigt die Dünndarmwandung eine sanduhrartige ...

human language representation



standardised annotations **SNOMED CT**

comparison



standardised annotations **SNOMED CT**

Manual annotations

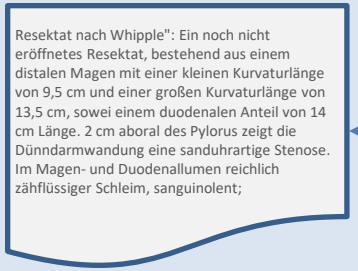
§ 7 Intact teeth are in the mouth.	11163003;245543004; 123851003
Fractures are visible on the medians of Mandible and Maxilla	263172003;263156006; 260528009
the fragments are dislocated.	123735002
Normal mucous membranes in mouth pharynx and on the larynx.	17621005;33044003; 71248005
Hyoid and thyroid cartilage are intact.	21387005;52940008; 11163003
Fragmental fractures of the two upper vertebrae of the cervical spine.	13321001;207984009; 207983003
Otherwise the cervical spine is intact.	122494005;11163003
Oesophagus as well as trachea are torn at the lower end of the neck.	262793000;282459005; 261122009;123958008

Nitroglycerin pump spray as required	387404004;385074009;225 761000
Amantadine bds	372763006;229799001
Allopurinol 300 ½ tablet every other day (last dose on 20091130)	387135004;385055001;225 760004
Mefenamic acid 500 mg up to 3x daily for pain in conjunction with	387185008;258684004; 229798009;22253000
simultaneous administration of a drug to protect the stomach e. g.	79970003;416118004; 373517009;69695003
Pantoprazole 40mg.	395821003;258684004
Torasemide bds	318034005;229799001
Melperone 50 mg p. m.	442519006;258684004; 422133006

ASSESS-CT: 60 documents annotated by clinical experts, 20 of them twice

http://assess-ct.eu/fileadmin/assess_ct/final_brochure/assessct_final_brochure.pdf

Annotation: Sources of complexity



map

SNOMED CT

Human language
- words, multiword terms
- syntactic structures
- relations at various levels

Ontology aspects
- classes
- relations
- logical constructors
- axioms

Terminology aspects
- concepts
- preferred terms
- synonyms
- text defs

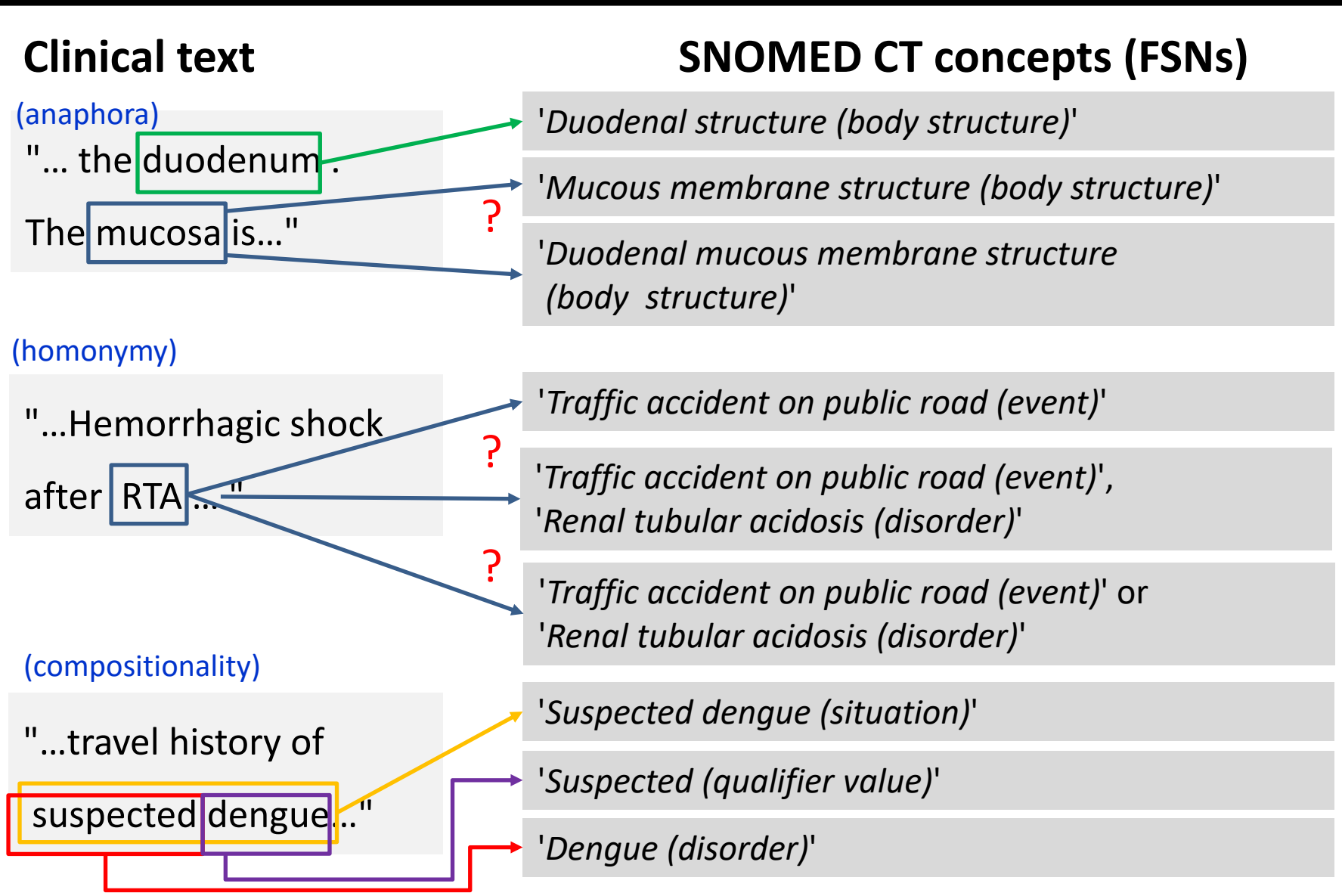
Clinical language
- Compact
- Un-grammatical
- Error-laden
- Context-dependent

Best text span to annotate?
Naïve or analytic annotation?

Known Issues

SNOMED CT
- Ill-defined concepts
- Similar concepts
- Pre-coordination vs. post-coordination
- Complex annotations (> 1 concept / term)

Examples


Clinical text
SNOMED CT concepts (FSNs)
(anaphora)
"… the duodenum .
The mucosa is…"
?
'Duodenal structure (body structure)'
'Mucous membrane structure (body structure)'
'Duodenal mucous membrane structure (body structure)'
(homonymy)
"…Hemorrhagic shock after RTA … "
?
?
?
'Traffic accident on public road (event)'
'Traffic accident on public road (event)', 'Renal tubular acidosis (disorder)'
'Traffic accident on public road (event)' or 'Renal tubular acidosis (disorder)'
(compositionality)
"…travel history of suspected dengue…"
'Suspected dengue (situation)'
'Suspected (qualifier value)'
'Dengue (disorder)'

Coding / Annotation guidelines

- Examples:
  1. German coding guidelines for ICD and OPS, 171 pages
  2. Using SNOMED CT in CDA models: 147 pages
  3. CHEMDNER-patents: annotation of chemical entities in patent corpus: annotation manual 30 pages
  4. CRAFT Concept Annotation guidelines: 47 pages
  5. Gene Ontology Annotation conventions: 7 pages
- Complex rule sets, requiring intensive training

1. http://www.dkgev.de/media/file/21502.Deutsche_Kodierrichtlinien_Version_2016.pdf
2. http://www.snomed.org/resource/resource/249
3. http://www.biocreative.org/media/store/files/2015/cemp_patent_guidelines_v1.pdf
4. http://bionlp-corpora.sourceforge.net/CRAFT/guidelines/CRAFT_concept_annotation_guidelines.pdf
5. http://geneontology.org/page/go-annotation-conventions

Focus of the talk

- What is semantic interoperability
- Analysis of manual SNOMED CT text annotations
- **Inter-annotator agreement (reliability)**
- Reasons for inter-annotator disagreement
- Discussion: how to improve agreement
- Conclusions / Outlook

- Assumption:
  better data reliability
  -> better semantic interoperability of clinical data

Principal quantitative results (English)

Concept coverage [95% CI]	SNOMED CT
Text annotations – English	.86 [.82-.88]

Term coverage [95% CI]	SNOMED CT
Text annotations – English	.68 [.64; .70]

Inter annotator agreement Krippendorff's Alpha [95% CI]	SNOMED CT
Text annotations	.37 [.33-.41]

(similar results with alternative annotation task, using non-SNOMED UMLS extract)

Krippendorff, Klaus (2013). Content analysis: An introduction to its methodology, 3rd edition. Thousand Oaks, CA: Sage.

Agreement map: SNOMED annotations

green: agreement – yellow: only annotated by one coder – red: disagreement - white no annotations

Focus of the talk

- What is semantic interoperability
- Analysis of manual SNOMED CT text annotations
- Inter-annotator agreement (reliability)
- **Reasons for inter-annotator disagreement**
- Discussion: how to improve agreement
- Conclusions / Outlook

- Assumption:
  better data reliability
  -> better semantic interoperability of clinical data

Systematic error analysis

- Creation of gold standard for SNOMED CT
  - 20 English text samples annotated twice →208 NPs
  - Analysis of English SNOMED CT annotations by two additional terminology experts
  - Consensus finding, according to pre-established annotation guidelines
- Inspection, analysis and classification of text annotation disagreements
- Presentation of some disagreement cases for SNOMED CT

Human issues

- Lack of domain knowledge / carelessness

Tokens	Annotator #1	Annotator #2	Gold standard
"IV"	'*Structure of abductor hallucis muscle (body structure)*'	'*Abducens nerve structure (body structure)* '	'*Abducens nerve structure (body structure)*'

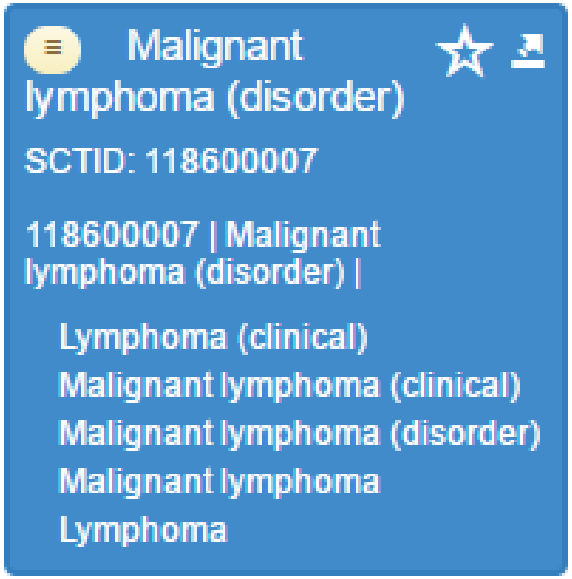
- Retrieval error (synonym not recognised)

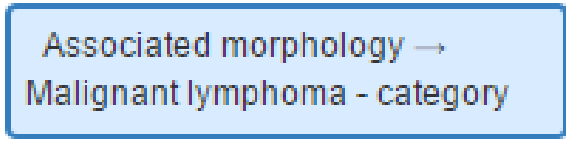
Tokens	Annotator #1	Annotator #2	Gold standard
"Glibenclamide"	'*Glyburide (substance)*'	–	'*Glyburide (substance)*'

Ontological issues (I)

- Logical polysemy ("dot categories")*

Tokens	Annotator #1	Annotator #2	Gold standard
'Lymphoma"	'*Malignant lymphoma (disorder)*'	'*Malignant lymphoma - category (morphologic abnormality)*'	'*Malignant lymphoma (disorder)*'



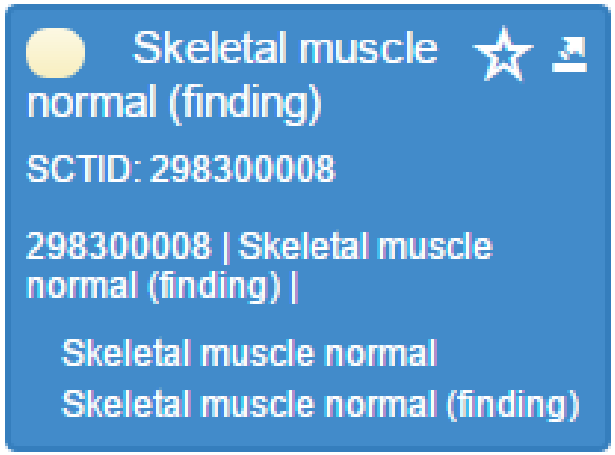


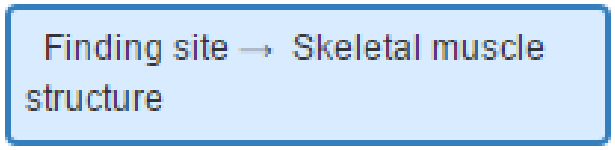
*Alexandra Arapinis, Laure Vieu: A plea for complex categories in ontologies. Applied Ontology 10(3-4): 285-296 (2015)

Ontological issues (II)

- Incomplete formal definitions

Tokens	Annotator #1	Annotator #2	Gold standard
"Motor: normal bulk and tone"	'*Skeletal muscle structure (body structure)*'	'*Muscle finding (finding)*'	'*Skeletal muscle normal (finding)*'
	'*Normal (qualifier value)*'	'*Normal (qualifier value)*'	





Ontological issues (II)

- Incomplete definitions

Tokens	Annotator #1	Annotator #2	Gold standard
"Motor: normal bulk and tone"	'*Skeletal muscle structure (body structure)*'	'*Muscle finding (finding)*'	'*Skeletal muscle normal (finding)*'
	'*Normal (qualifier value)*'	'*Normal (qualifier value)*'	

Tokens	Annotator #1	Annotator #2	Gold standard
"Former smoker"	'*In the past (qualifier value)*'	'*History of (contextual qualifier) (qualifier value)*'	'*Ex-smoker (finding)*'
	'*Smoker (finding)*'	'*Smoker (finding)*'	

Ontological issues (III)

- Navigational concepts vs. observables

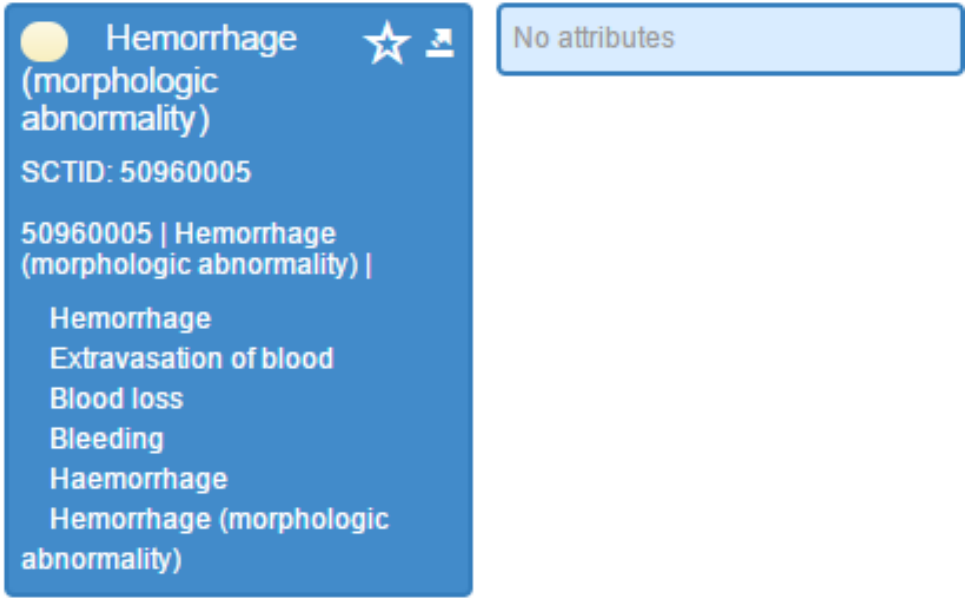
Tokens	Annotator #1	Annotator #2	Gold standard
"palpebral fissure"	*Finding of measures of palpebral fissure (finding)*	*Structure of palpebral fissure (body structure)*	*Measure of palpebral fissure (observable entity)*

- Fuzzy, undefined qualifiers

Tokens	Annotator #1	Annotator #2	Gold standard
"Significant bleeding"	'*Significant (qualifier value)*'	'*Severe (severity modifier) (qualifier value)*'	'*Moderate (severity modifier) (qualifier value)*'
	'*Bleeding (finding)*'	'*Bleeding (finding)*'	'*Bleeding (finding)*'

Interface term (synonym) issues

Tokens	Annotator #1	Annotator #2	Gold standard
"Blood extra-vacation"	'*Blood (substance)*'	'*Hemorrhage (morphologic abnormality)*'	*'Hemorrhage (morphologic abnormality)'*
	'*Extravasation (morphologic abnormality)*'	"extravasation of blood"	



Interface term (synonym) issues

Tokens	Annotator #1	Annotator #2	Gold standard
"Blood extra-vasation"	'*Blood (substance)*' '*Extravasation (morphologic abnormality)*'	'*Hemorrhage (morphologic abnormality)*'	*'Hemorrhage (morphologic abnormality)'*
		"extravasation of blood"	

Tokens	Annotator #1	Annotator #2	Gold standard
"anxious"	'*Anxiety (finding)*'	'*Worried (finding)*'	'*Anxiety (finding)*'
		"anxious cognitions"	

Focus of the talk

- What is semantic interoperability
- Analysis of manual SNOMED CT text annotations
- Inter-annotator agreement (reliability)
- Reasons for inter-annotator disagreement
- **Discussion: how to improve agreement**
- Conclusions / Outlook

- Assumption:
  better data reliability
  -> better semantic interoperability of clinical data

Prevention and remediation of annotation disagreements

- Rationales:
  - More principled SNOMED CT coding of EHR content
  - More principled binding of SNOMED CT codes to clinical models
  - Consistent manual annotations for training corpora and reference standards
  - Improvement of performance of NLP-based annotations

Prevention and remediation of annotation disagreements

Prevention: annotation processes

- Training with continuous feedback
  - Early detection of inter annotator disagreement triggers guideline enforcement / revision
- Tooling
  - Optimised concept retrieval (fuzzy, substring, synonyms)
  - Guideline enforcement by appropriate tools
  - Postcoordination support (complex syntactic expressions instead of simple concept grouping)

Prevention: improve SNOMED CT quality

- Fill gaps
  - Add missing equivalence axioms
  - Self-explaining labels, text definitions where necessary
- Preference rules to manage polysemy
- Strengthen ontological foundations
  - Upper-level ontology alignment
  - Better distinction between domain entities and information entities
  - Overhaul problematic subhierarchies, especially qualifiers

Prevention: improve content maintenance

- Data-driven terminology maintenance
  - Harvest notorious disagreements between annotations from clinical datasets
  - Detect imbalances by analysing concept frequency and co-occurrence between comparable institutions
  - Community processes: crowdsourcing of interface terms by languages, dialects, specialties, user groups (*ASSESS-CT: interface terminologies to be maintained separately from reference terminologies)*

Remediation of annotation disagreements

Remediation of annotation disagreements

- Exploit ontological dependencies / implications

Concept A	Concept B	Dependency
'*Mast cell neoplasm (disorder)*'	'*Mast cell neoplasm (morphologic abnormality)*'	*A* subclassOf **AssociatedMorphology** some *B*
'*Isosorbide dinitrate' (product)*'	'*Isosorbide dinitrate (substance)*'	*A* subclassOf **HasActiveIngredient** some *B*
'*Palpation (procedure)*'	'*Palpation - action (qualifier value)*'	*A* subclassOf **Method** some *B*
'*Blood pressure taking (procedure)*'	'*Blood pressure (observable entity)*'	*A* subclassOf **hasOutcome** some *B*
'*Increased size (finding)*'	'*Increased (qualifier value)*'	*A* subclassOf **isBearerOf** some *B*
'*Finding of heart rate (finding)*'	'*Heart rate (observable entity)*'	*A* subclassOf **Interprets** some *B*

Experiment

- Gold standard expansion:
  - Step 1: include concepts linked by attributive relations:
    - *A* subclassOf **Rel** some *B*
  - Step 2: include additional first-level taxonomic relations:
    - *A* subclassOf *B*

Language of text sample	Gold standard expansion	F measure
English	no expansion	0.28
	expansion step 1	0.28
	expansion step 2	0.29

- only insignificant improvement
- possibly due to missing relations in SNOMED CT (see "former smoker" and "skeletal muscle normal" examples)
- just a side issue… requires more investigation

Focus of the talk

- What is semantic interoperability
- Analysis of manual SNOMED CT text annotations
- Inter-annotator agreement (reliability)
- Reasons for inter-annotator disagreement
- Discussion: how to improve agreement
- **Conclusions / Outlook**

  - Assumption:
    better data reliability
    -> better semantic interoperability of clinical data

Conclusions

- Poor agreement hampers SNOMED CT use:
  - Clinical decision support, cohort building, content retrieval, summarisation, analytics,…
    (but not specific for SNOMED CT → ACCESS CT)
- Prevention & Remediation:
  - Education, tooling, guidelines
  - Large-scale SNOMED CT content and structure improvement
  - High coverage local *interface terminologies*, representing real language of clinicians

Outlook

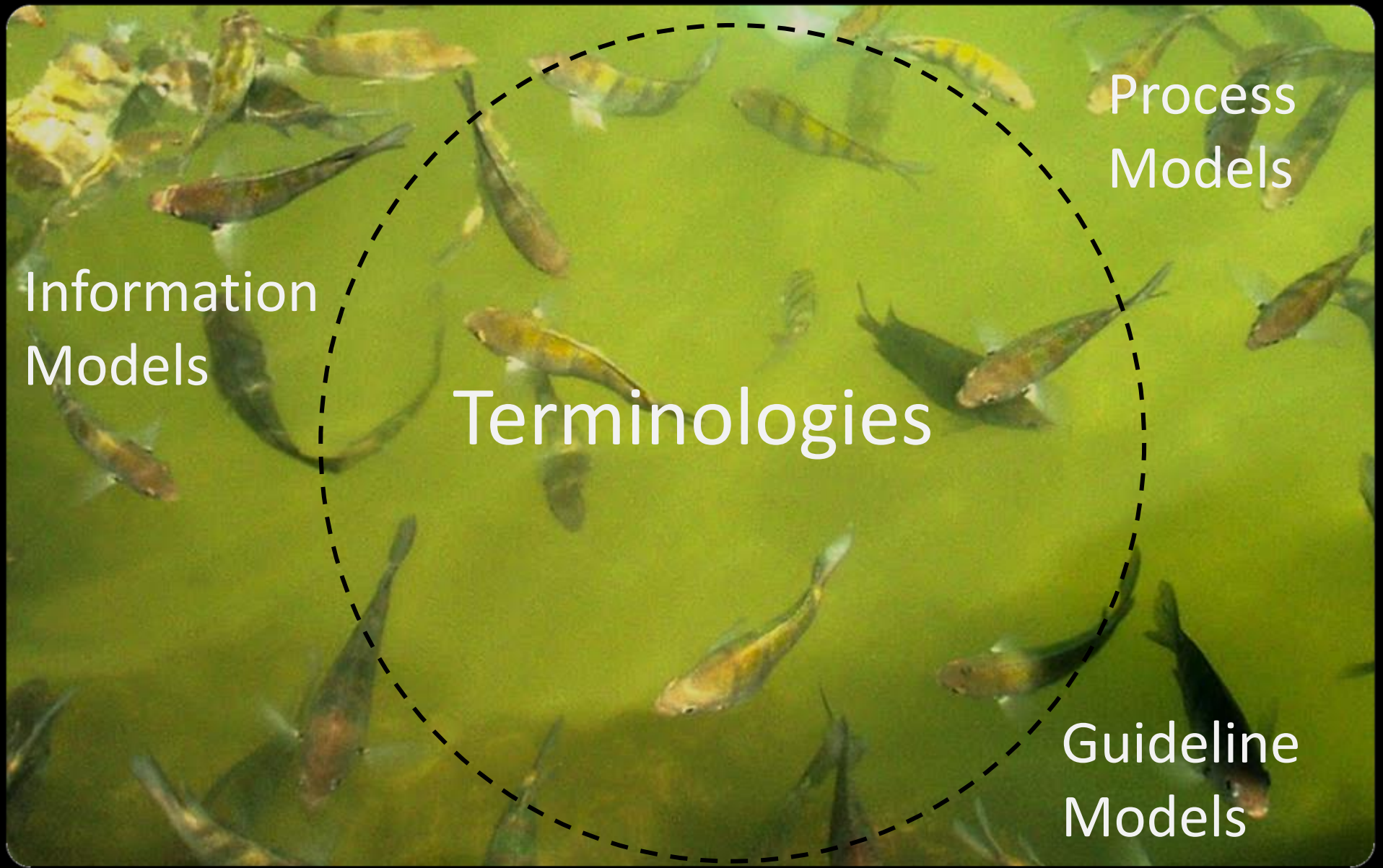
- "Learning systems" for improvement terminology content / structure / tooling.
- "Clinical big data": pooling of non-re-identifiable annotations from multiple institutions
- Text preprocessing: resolution of ambiguous short forms
- Post processing of SNOMED CT annotations: Stream of codes → text knowledge graph
- Community efforts for interface terminology creation and maintenance

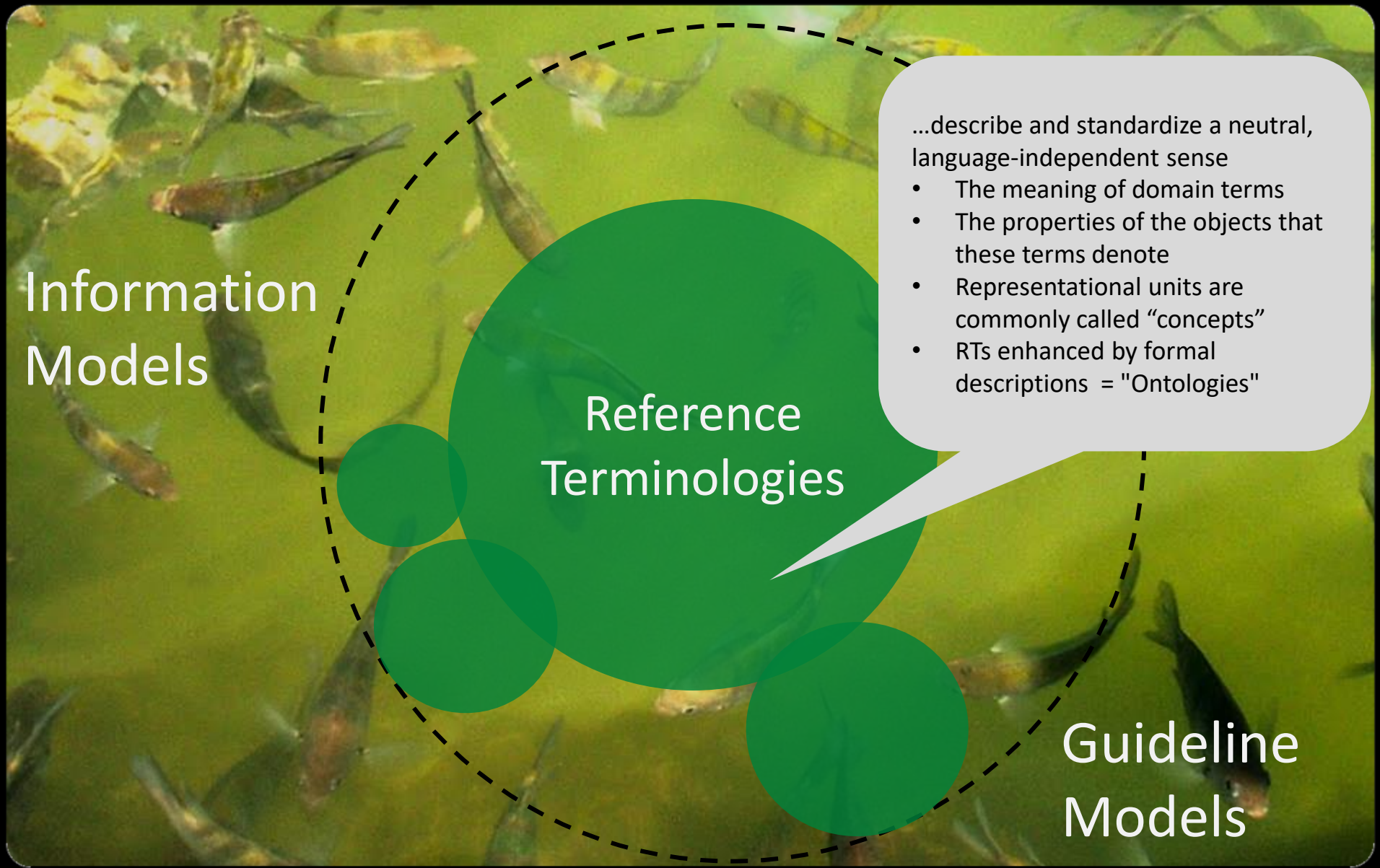
*Martínez-Costa, Kalra, Schulz. Semantic enrichment of clinical models towards semantic interoperability. JAMIA 2015 May;22(3):565-76

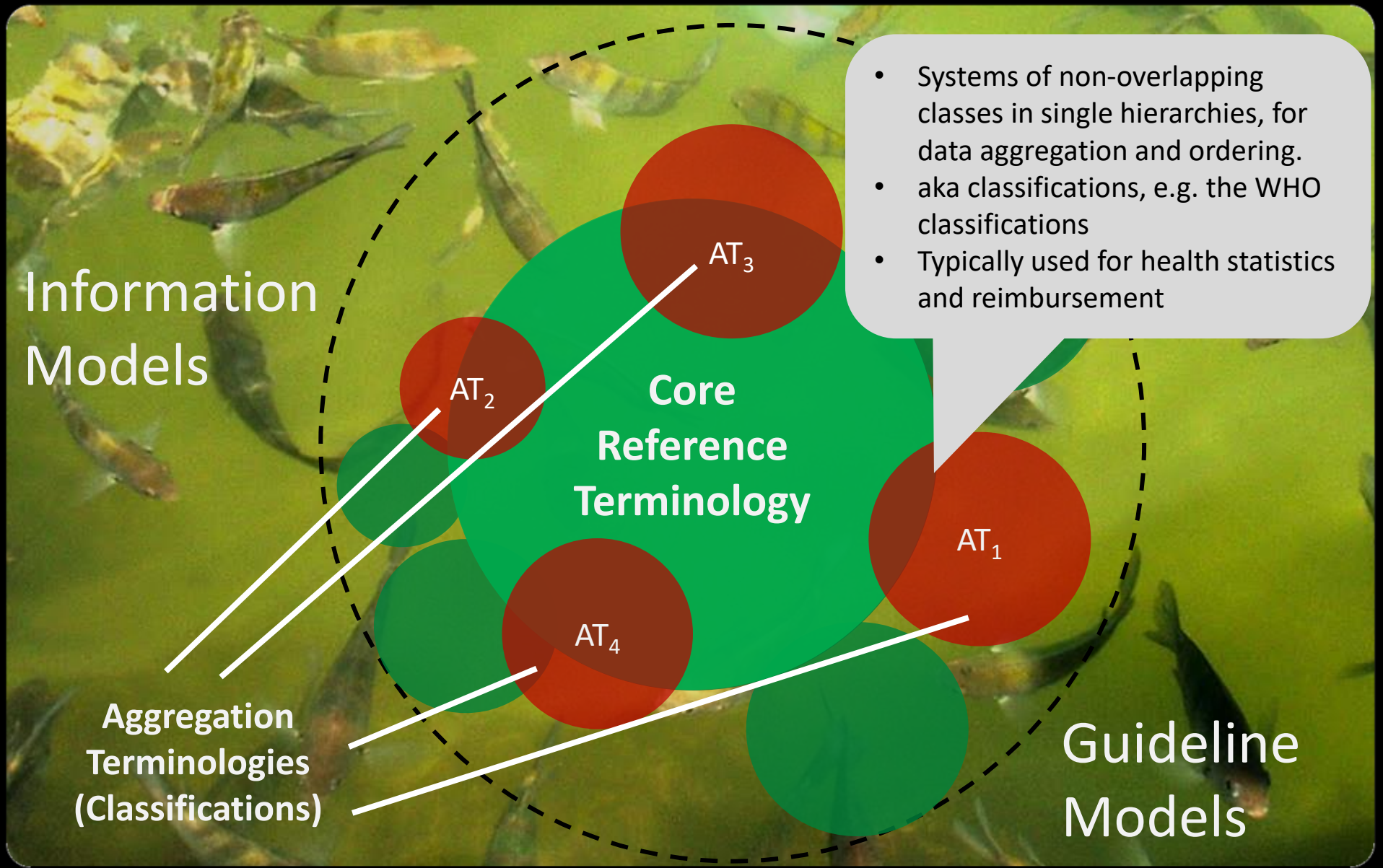
Thanks for your attention

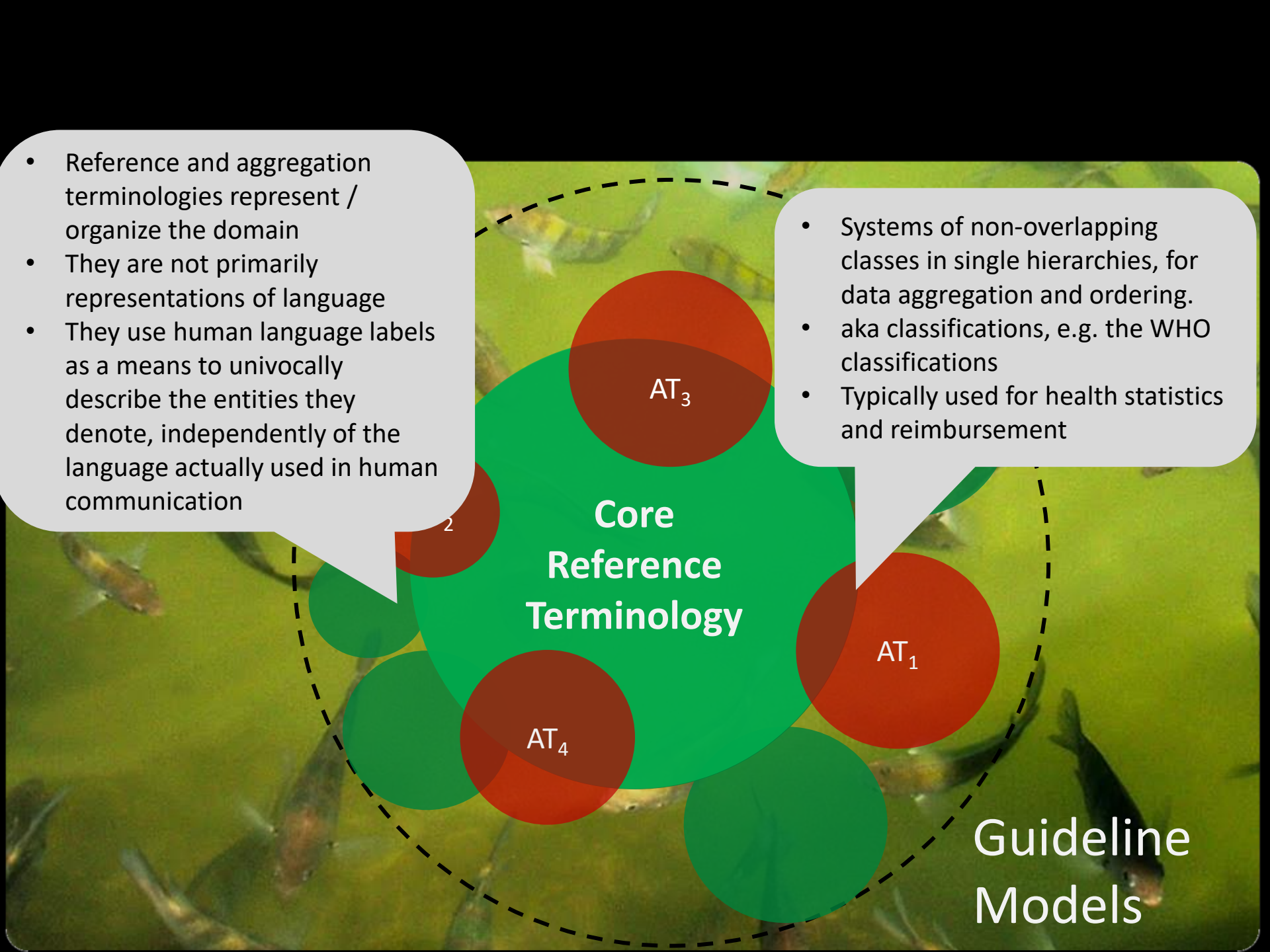
- Slides will be made accessible at purl.org/steschu
- Acknowledgements: ASSESS CT team:
Jose Antonio Miñarro-Giménez, Catalina Martínez-Costa, Daniel Karlsson, Kirstine Rosenbeck Gøeg, Kornél Markó, Benny Van Bruwaene, Ronald Cornet, Marie-Christine Jaulent, Päivi Hämäläinen, Heike Dewenter, Reza Fathollah Nejad, Sylvia Thun, Veli Stroetmann, Dipak Kalra
- Contact: stefan.schulz@medunigraz.at

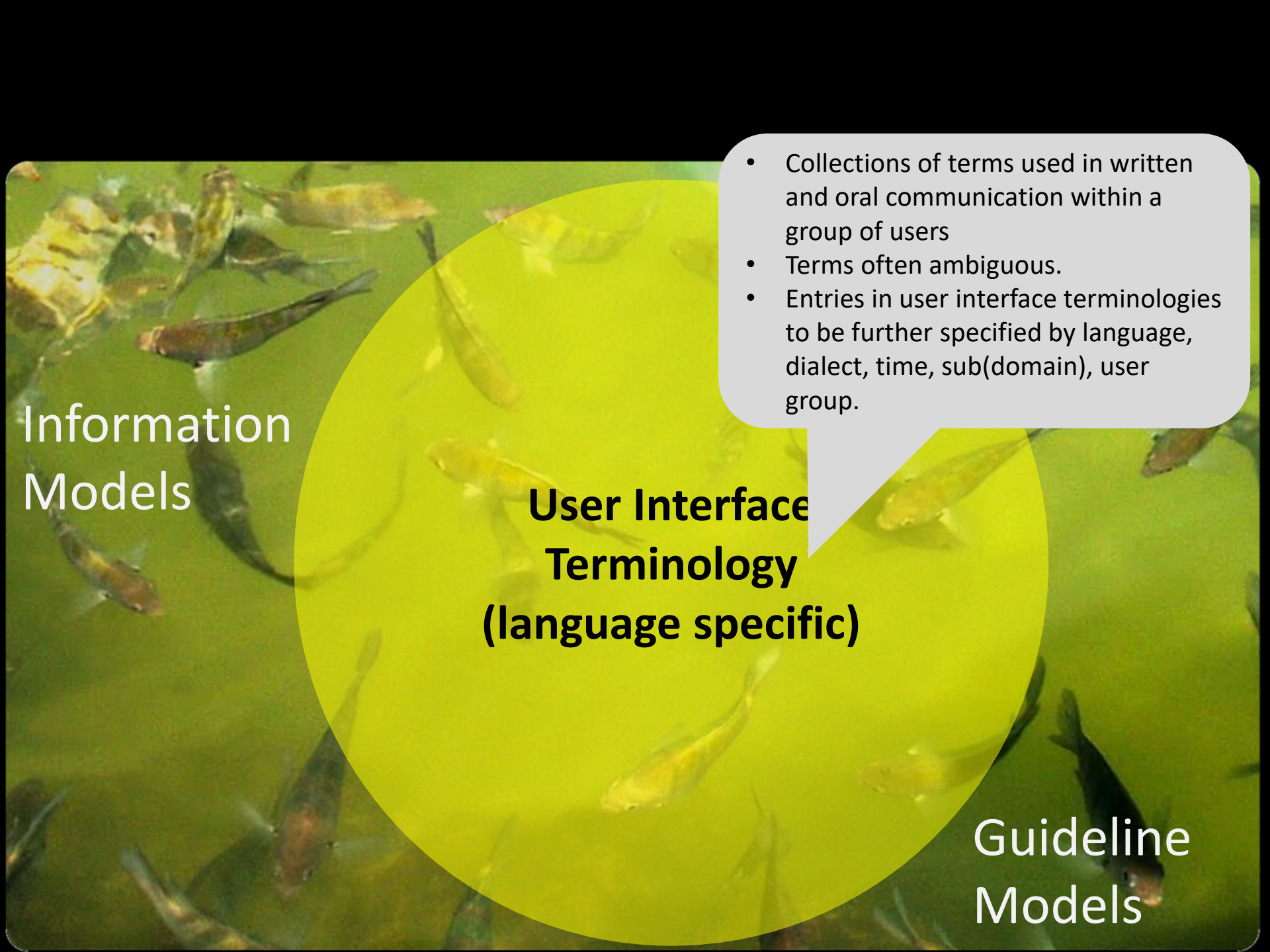
Ecosystem of semantic assets

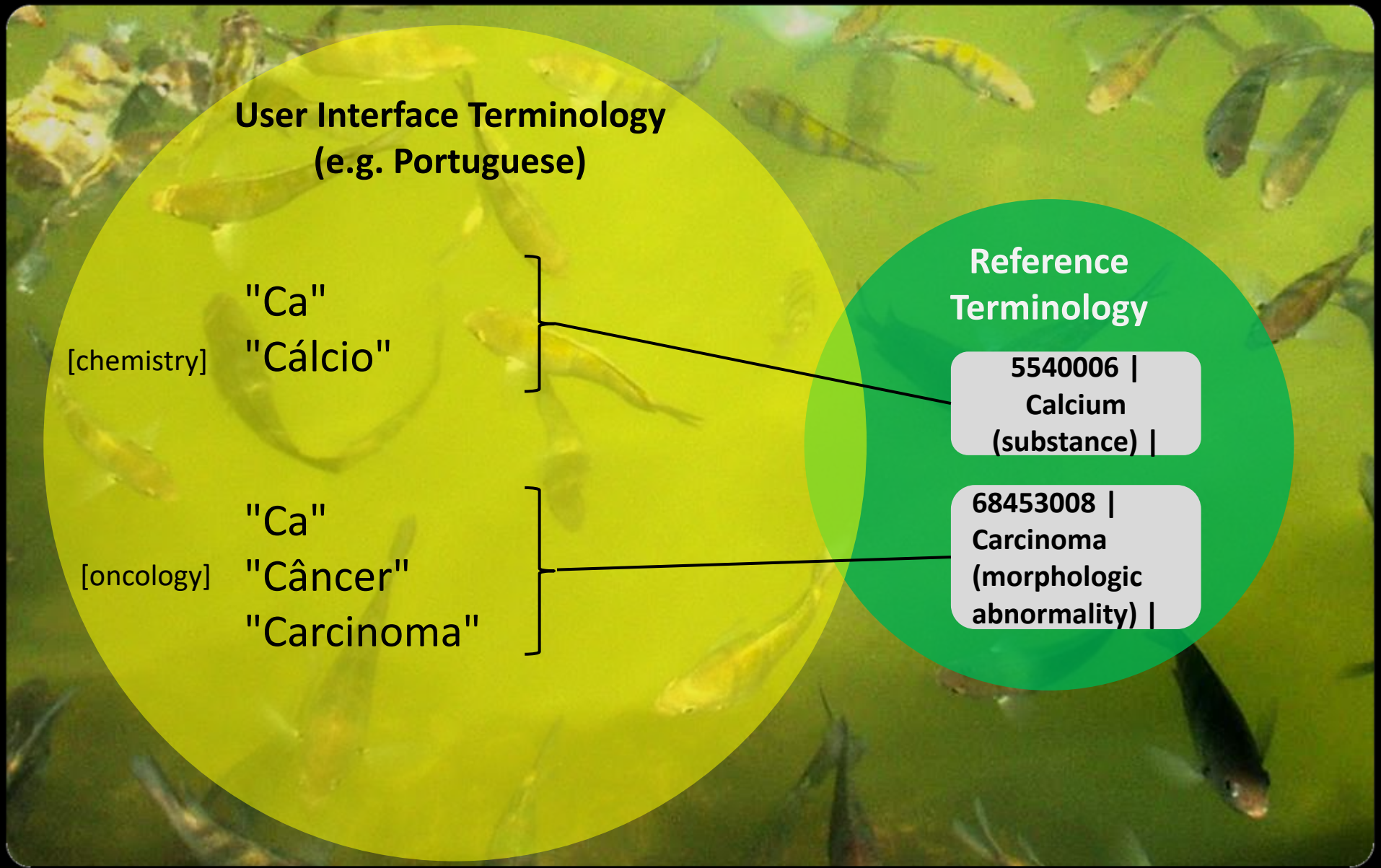


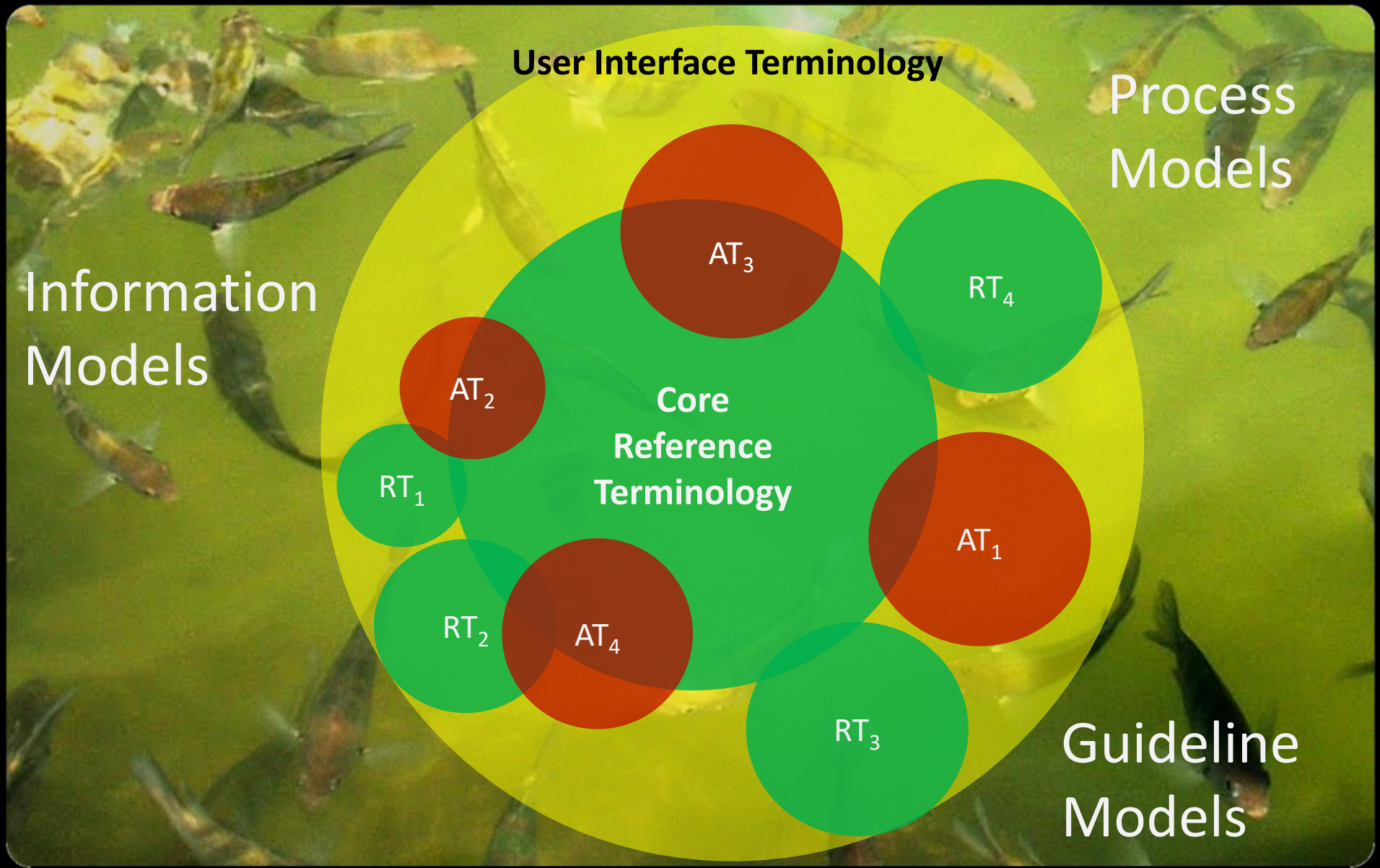

Information Models
Reference Terminologies
Guideline Models
…describe and standardize a neutral, language-independent sense
• The meaning of domain terms
• The properties of the objects that these terms denote
• Representational units are commonly called “concepts”
• RTs enhanced by formal descriptions = "Ontologies"


Information Models
AT3
AT2
Core Reference Terminology
AT1
AT4
Aggregation Terminologies (Classifications)
Guideline Models
• Systems of non-overlapping classes in single hierarchies, for data aggregation and ordering.
• aka classifications, e.g. the WHO classifications
• Typically used for health statistics and reimbursement

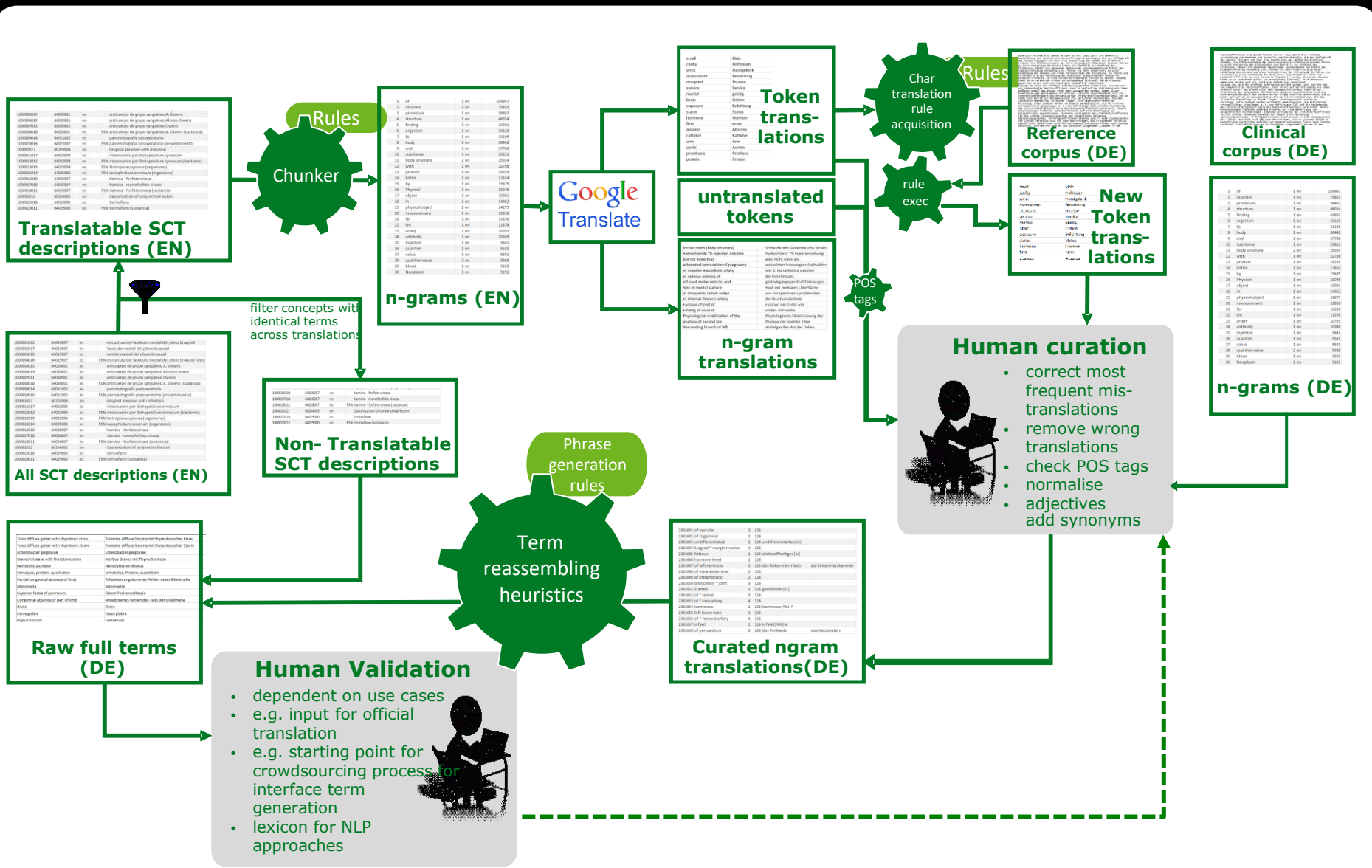

Reference and aggregation terminologies represent / organize the domain
They are not primarily representations of language
They use human language labels as a means to univocally describe the entities they denote, independently of the language actually used in human communication
Systems of non-overlapping classes in single hierarchies, for data aggregation and ordering.
aka classifications, e.g. the WHO classifications
Typically used for health statistics and reimbursement
AT3
AT2
AT1
AT4
Core Reference Terminology
Guideline Models


Information Models
User Interface Terminology (language specific)
Guideline Models
• Collections of terms used in written and oral communication within a group of users
• Terms often ambiguous.
• Entries in user interface terminologies to be further specified by language, dialect, time, sub(domain), user group.


User Interface Terminology
(e.g. Portuguese)
[chemistry]
"Ca"
"Cálcio"
[oncology]
"Ca"
"Câncer"
"Carcinoma"
Reference Terminology
5540006 | Calcium (substance) |
68453008 | Carcinoma (morphologic abnormality) |


User Interface Terminology
Process Models
Information Models
Guideline Models
Core Reference Terminology
AT3
RT4
AT2
RT1
AT1
RT2
AT4
RT3

MUG-GIT: Creation of German Interface Terminologie for SNOMED CT



ngram – core vocabulary

vaginal	1	1478	vaginales|JJ	Scheiden-	
fluoroscopic guidance	2	1477	Durchleuchtungskontrolle|NN|F		
disc	1	1476	Scheibe|NN|F		
lower limb	2	1473	unteres|JJ Extremität|NN|F	Bein|NN|N	
brain	1	1468	Gehirn|NN|N	Hirn|NN|N	Encephalon|NN|N
preparation	1	1464	Zubereitung|NN|F	Aufbereitung|NN|F	Präparation|NN|F
method	1	1463	Verfahren|NN|N	Methode|NN|F	
of bone	2	1462	des Knochens	_Knochen_	
Red	1	1455	rotes|JJ		
Monitoring	1	1453	Überwachung|NN|F	Monitoring|NN|N	
Computed	1	1453	berechnetes|JJ	Computer-	
phalanx	1	1449	Phalanx|NN|F		
subsp.	1	1449			
anastomosis	1	1447	Anastomose|NN|F	Anastomosierung|NN|F	
vessel	1	1446	Blutgefäß|NN|N	Gefäß|NN|N	
Computed tomography	2	1443	Computertomographie|NN|F		
uterus	1	1436	Uterus|NN|M	Gebärmutter|NN|F	
difficulty	1	1432	Schwierigkeit|NN|F		
elbow	1	1429	Ellbogen|NN|M	Cubitus|NN|M	Ellbogengelenk|NN|N
high	1	1429	hohes|JJ		
food	1	1423	Lebensmittel|NN|N	Speise|NN|F	Nahrungsmittel|NN|N
Observation	1	1423	Beobachtung|NN|F		
using fluoroscopic	2	1422			
unable	1	1421	unfähiges|JJ		
Peripheral	1	1419	peripheres|JJ		
unable to	2	1418	unfähig zu		
Vascular	1	1417	vaskuläres|JJ	Gefäß-	
using fluoroscopic guidance	3	1416	mit Durchleuchtungskontrolle		
Benign neoplasm	2	1415	gutartiges|JJ Neubildung|NN|F	gutartiges|JJ Neoplasie|NN|F	benignes|JJ Neoplasie|NN|F

Machine-generated Interface terms

20170315_240011_002	126952004	Neoplasm of brain	Gehirnneubildung
20170315_240011_003	126952004	Neoplasm of brain	Neubildung des Hirns
20170315_240011_004	126952004	Neoplasm of brain	Hirnneubildung
20170315_240011_005	126952004	Neoplasm of brain	Neoplasie des Gehirns
20170315_240011_006	126952004	Neoplasm of brain	Gehirnneoplasie
20170315_240011_007	126952004	Neoplasm of brain	Neoplasie des Hirns
20170315_240011_008	126952004	Neoplasm of brain	Hirnneoplasie
20170315_240011_009	126952004	Neoplasm of brain	Neoplasma des Gehirns
20170315_240011_010	126952004	Neoplasm of brain	Gehirnneoplasma
20170315_240011_011	126952004	Neoplasm of brain	Neoplasma des Hirns
20170315_240011_012	126952004	Neoplasm of brain	Hirnneoplasma
20170315_241010_001	126953009	Neoplasm of cerebrum	Neubildung des Großhirns
20170315_241010_002	126953009	Neoplasm of cerebrum	Neoplasie des Großhirns
20170315_241010_003	126953009	Neoplasm of cerebrum	Neoplasma des Großhirns
20170315_242015_001	126954003	Neoplasm of frontal lobe	Neubildung des Frontallappens
20170315_242015_002	126954003	Neoplasm of frontal lobe	Neubildung des Lobus frontalis
20170315_242015_003	126954003	Neoplasm of frontal lobe	Neoplasie des Frontallappens
20170315_242015_004	126954003	Neoplasm of frontal lobe	Neoplasie des Lobus frontalis
20170315_242015_005	126954003	Neoplasm of frontal lobe	Neoplasma des Frontallappens
20170315_242015_006	126954003	Neoplasm of frontal lobe	Neoplasma des Lobus frontalis
20170315_243013_001	126955002	Neoplasm of temporal lobe	Neubildung des Temporallappens
20170315_243013_002	126955002	Neoplasm of temporal lobe	Neubildung des Lobus temporalis
20170315_243013_003	126955002	Neoplasm of temporal lobe	Neoplasie des Temporallappens
20170315_243013_004	126955002	Neoplasm of temporal lobe	Neoplasie des Lobus temporalis
20170315_243013_005	126955002	Neoplasm of temporal lobe	Neoplasma des Temporallappens