MIRACUM Kick-off Symposium
22nd - 23rd February 2018
Erlangen

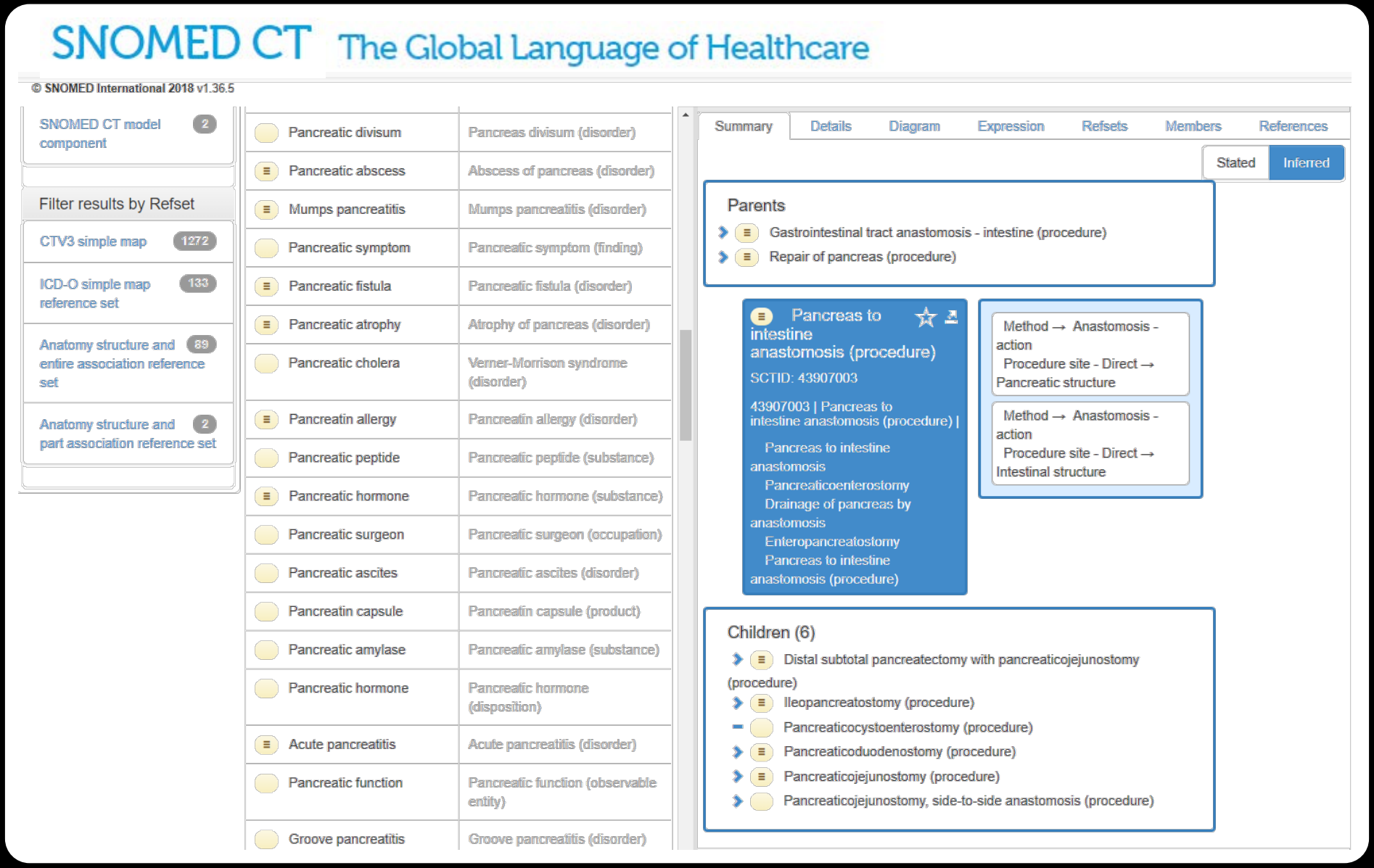
miracum
Medical Informatics for Research and Care in University Medicine

Stefan Schulz
Medical University of Graz (Austria)

purl.org/steschu

The use of **SNOMED CT** in BMBF MI-I: Opportunities, Challenges and Requirements

SNOMED CT – An opportunity for MI-I ?



SNOMED CT – An opportunity for MI-I ?

- Covers all aspects of the EHR, providing metadata ("concepts") for semantic interoperability of clinical data
- ~300k clinical concepts (disorders, findings, procedures, observables, organisms, products, substances, qualities, …)
- Terms in several languages
  (no official German translation)
- Increasing world-wide acceptance
- Computable and Extensible
- International standard, governed by an non-profit SDO ("SNOMED International", former IHTSDO)

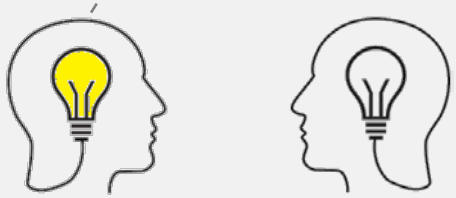
SNOMED CT – Yet another standard?

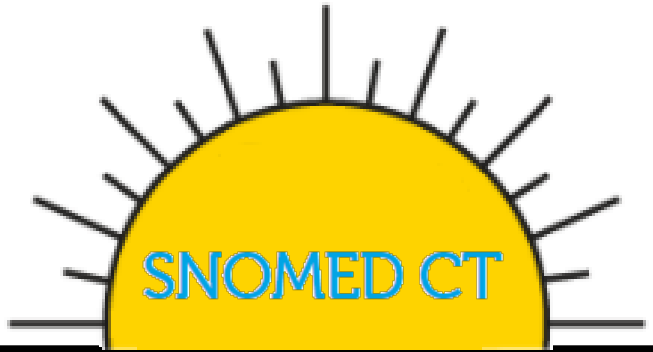
current

There are **n** competing semantic standards

HL7 NCIT ICD LOINC
IDMP ICPC DICOM HPO
RxNORM MedDRA ADT
MeSH OPS Abdamed

"**n?** Ridiculous ! Let's develop one global, universal standard that covers everybody's use cases"



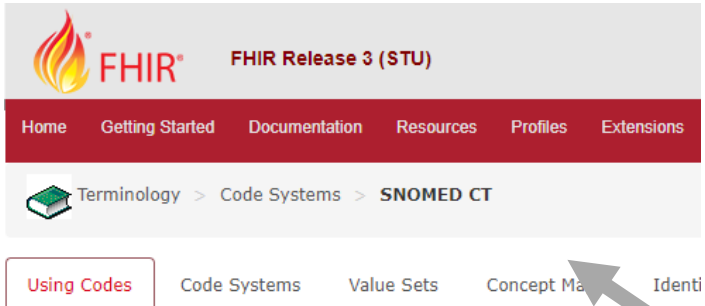


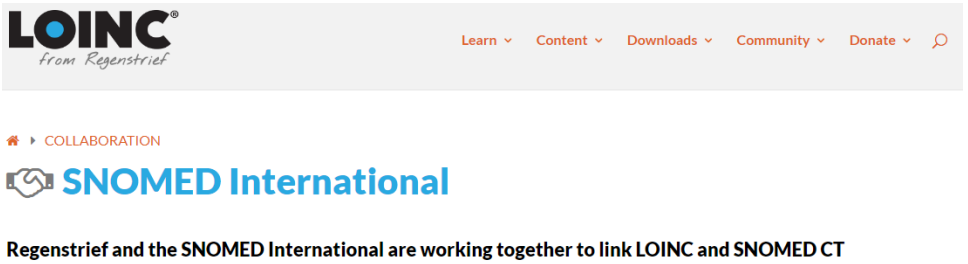
soon

There are **n + 1** competing semantic standards

HL7 NCIT ICD ICPC HPO
IDMP DICOM SNOMED CT
RxNORM MedDRA ADT LOINC
MeSH OPS Abdamed

Designed to interface with other standards





SNOMED CT

SNOMED International
Leading healthcare terminology, worldwide

IDMP
Identification of Medicinal Products
Data elements and structures for the unique identification and exchange

4. Short Term (July 2017, Jan 2018 International Release)

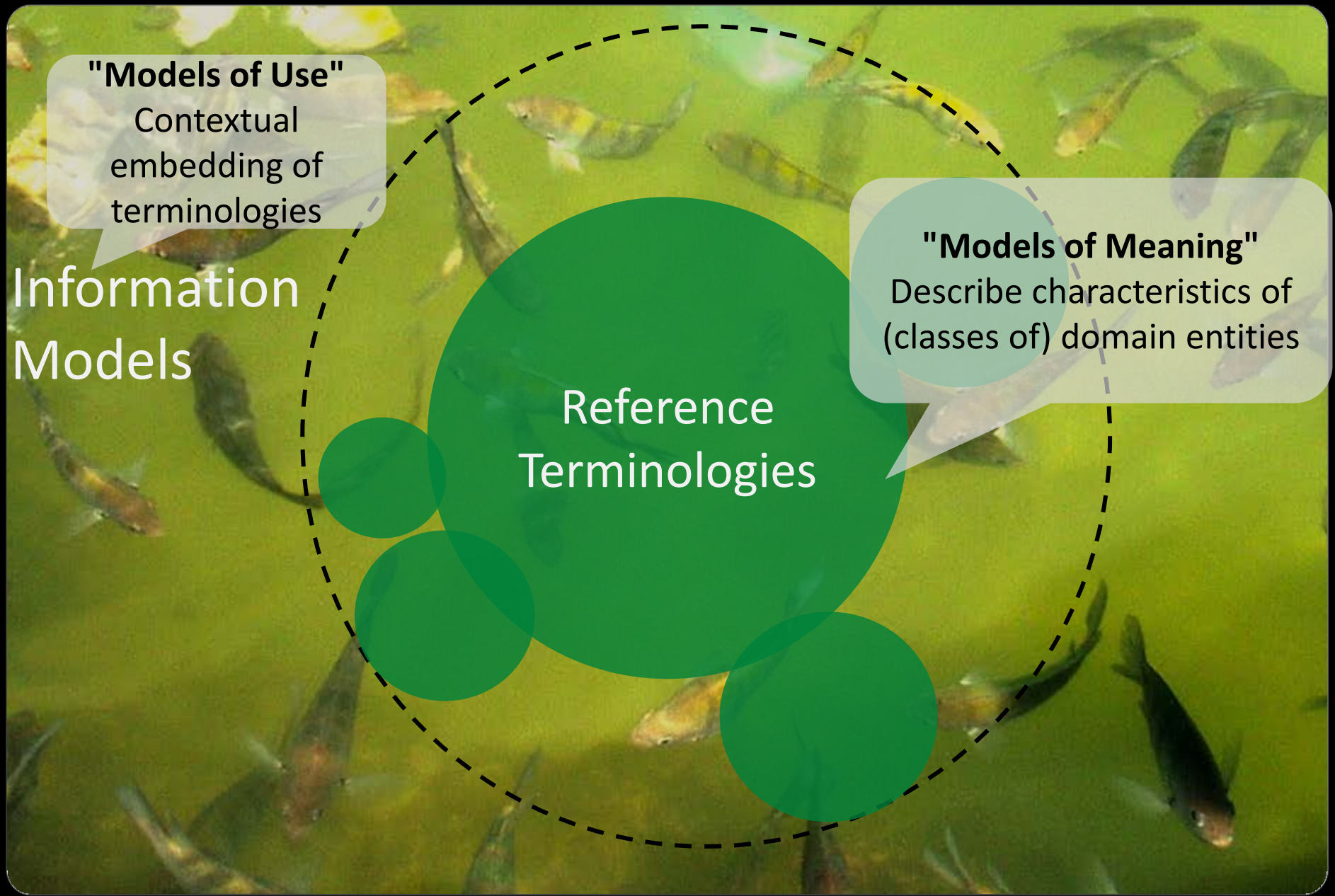
   The short term goals of the SNOMED Drug project will include an agile, risk adverse approach in the area of drug products specifically to;

4.1. Provide coarse level content internationally, supporting granularity at national level.
4.2. Focus on international requirements and regulations, such as initial IDMP alignment to support interoperability defined by the use cases and remove some barriers of adoption.

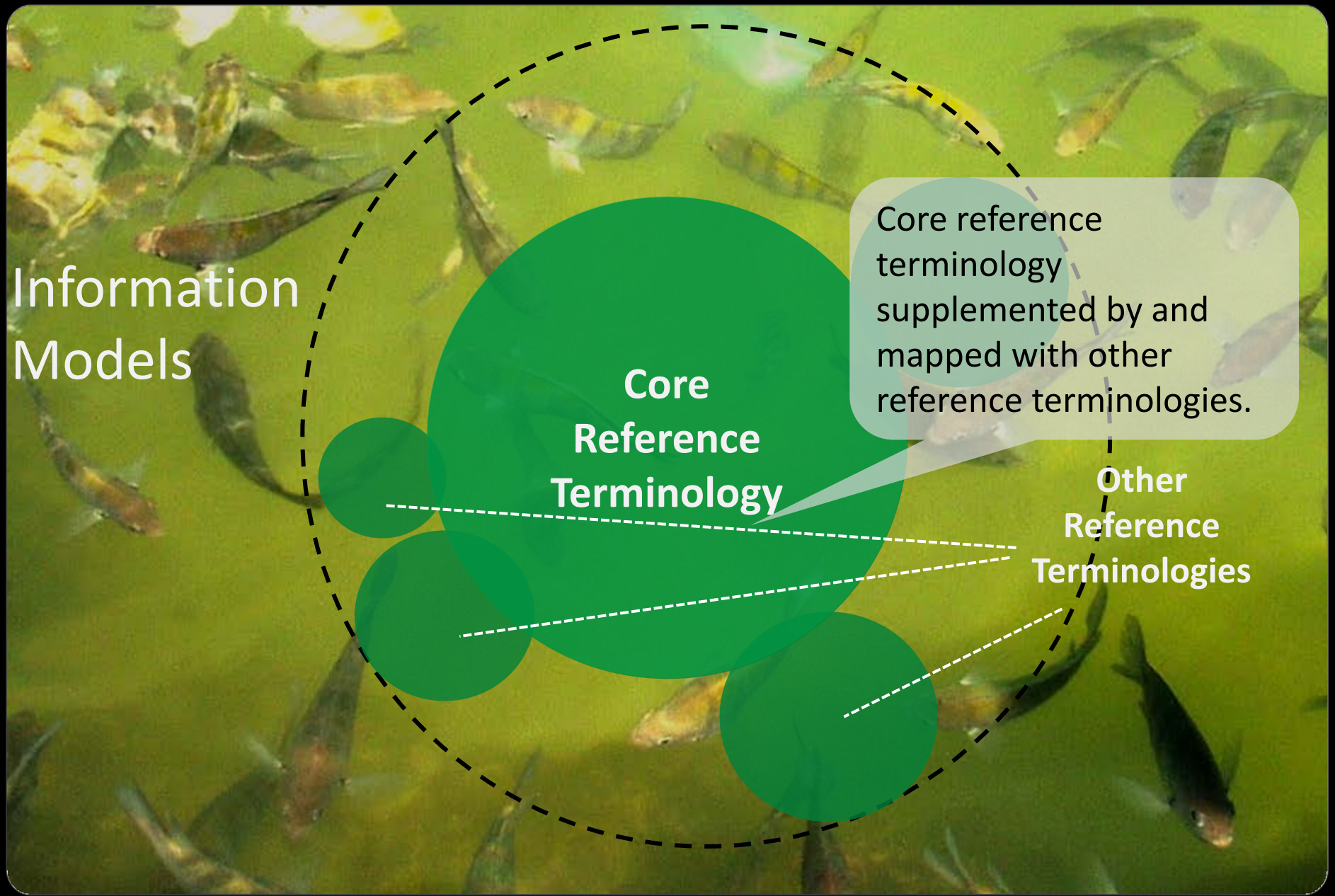
INTERNATIONAL HEALTH TERMINOLOGY STANDARDS DEVELOPMENT ORGANISATION
SNOMED CT®
G•M•D•N

GMDN and IHTSDO sign Cooperation Agreement

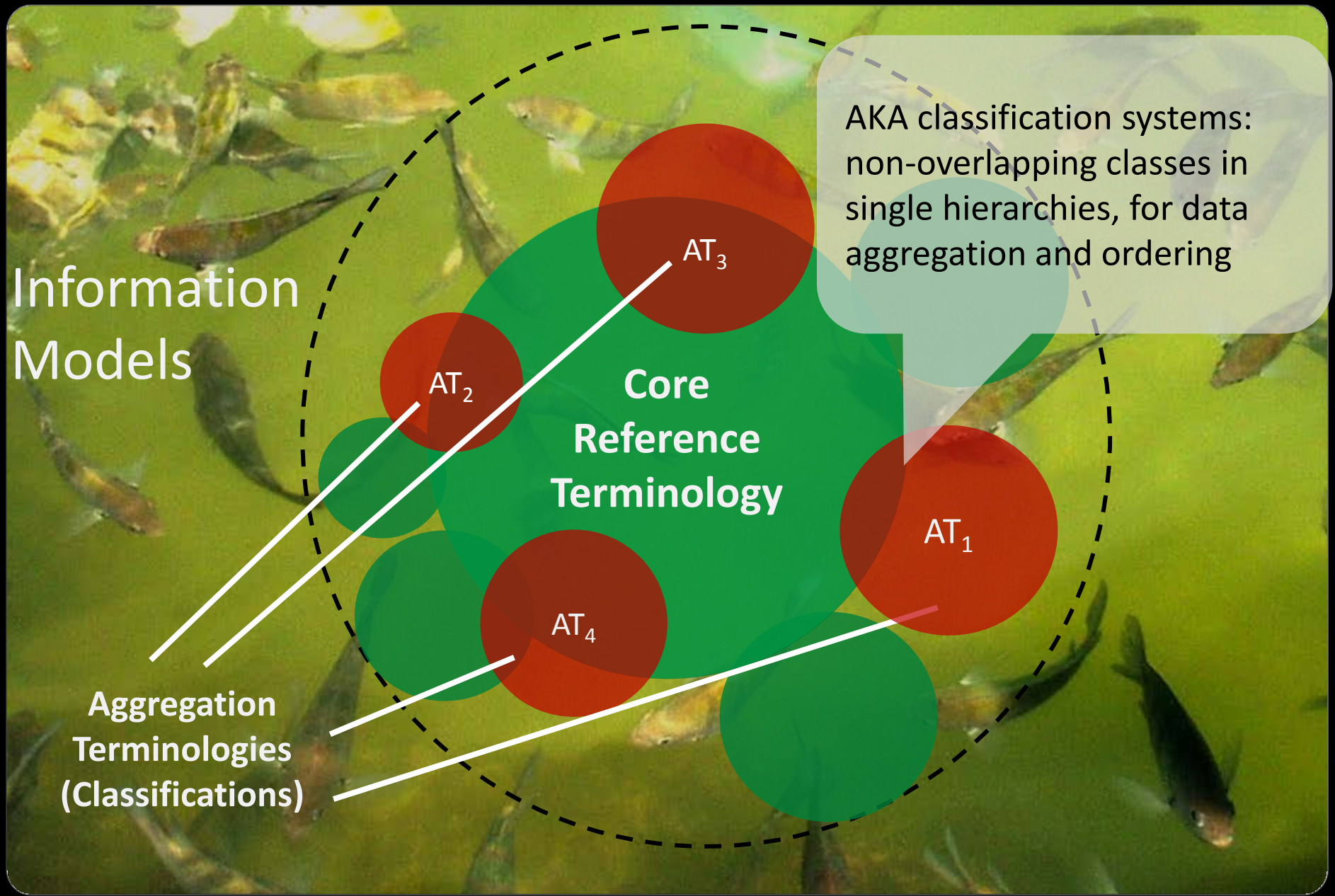
Interoperability ecosystem



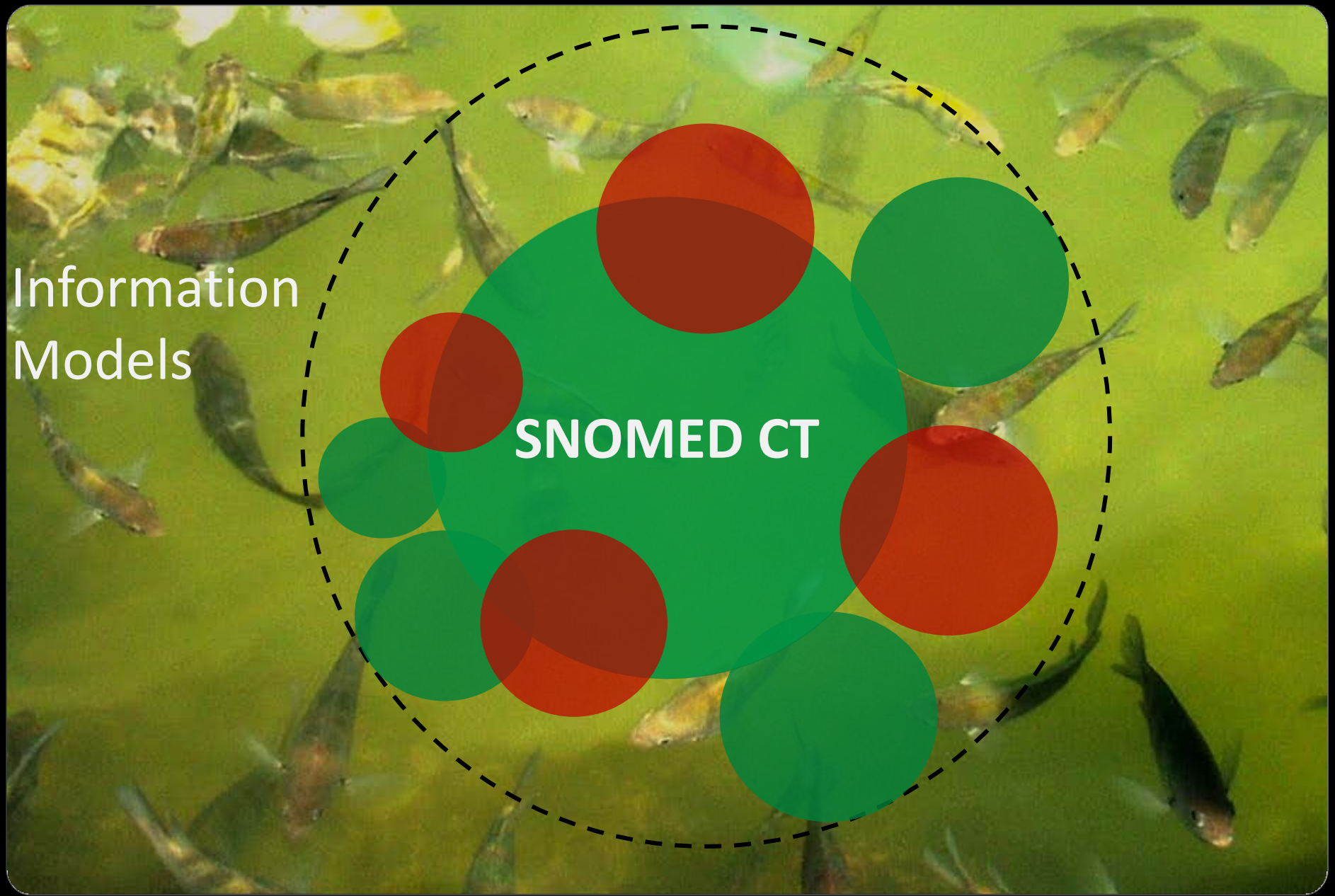
Interoperability ecosystem



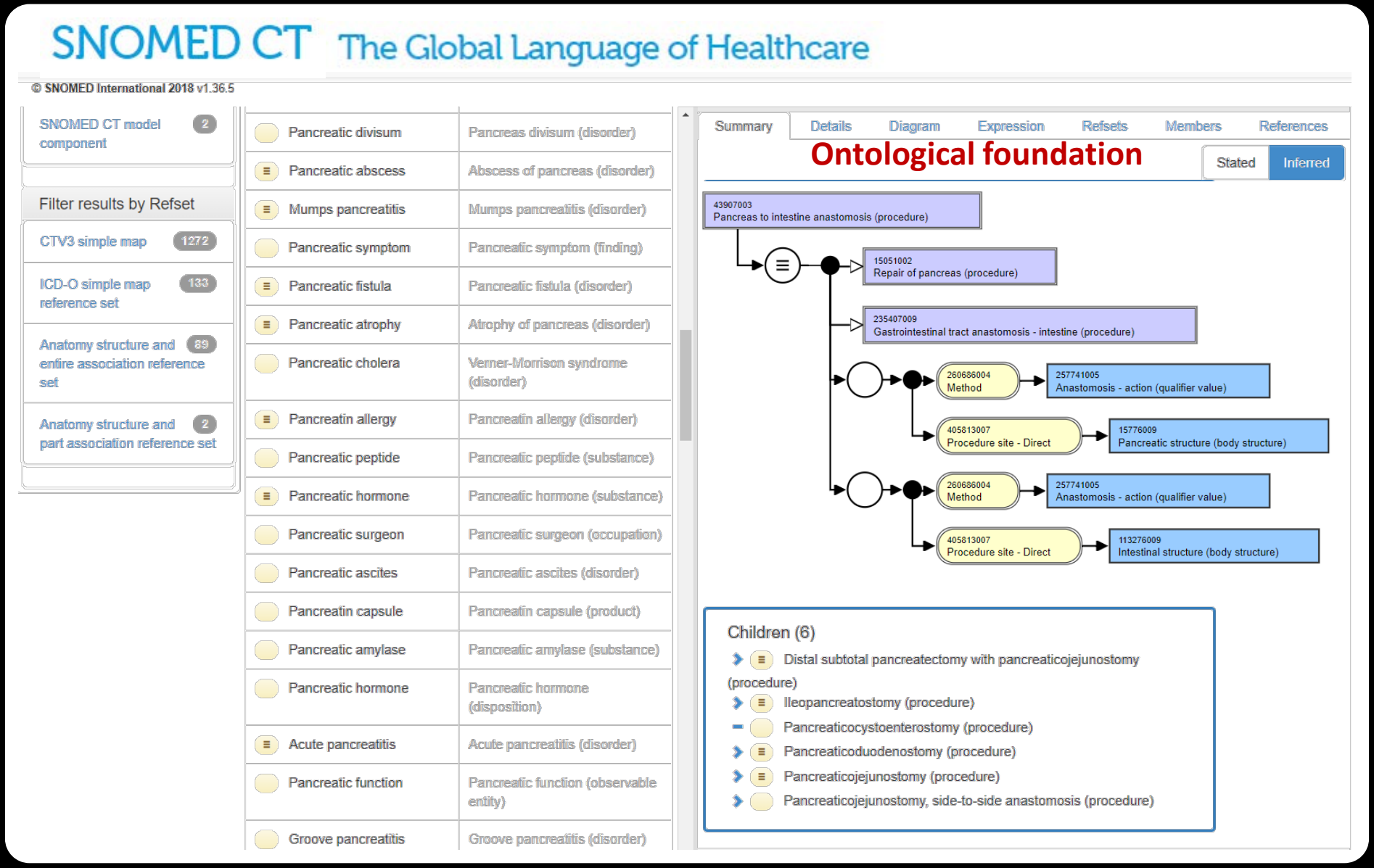
Interoperability ecosystem



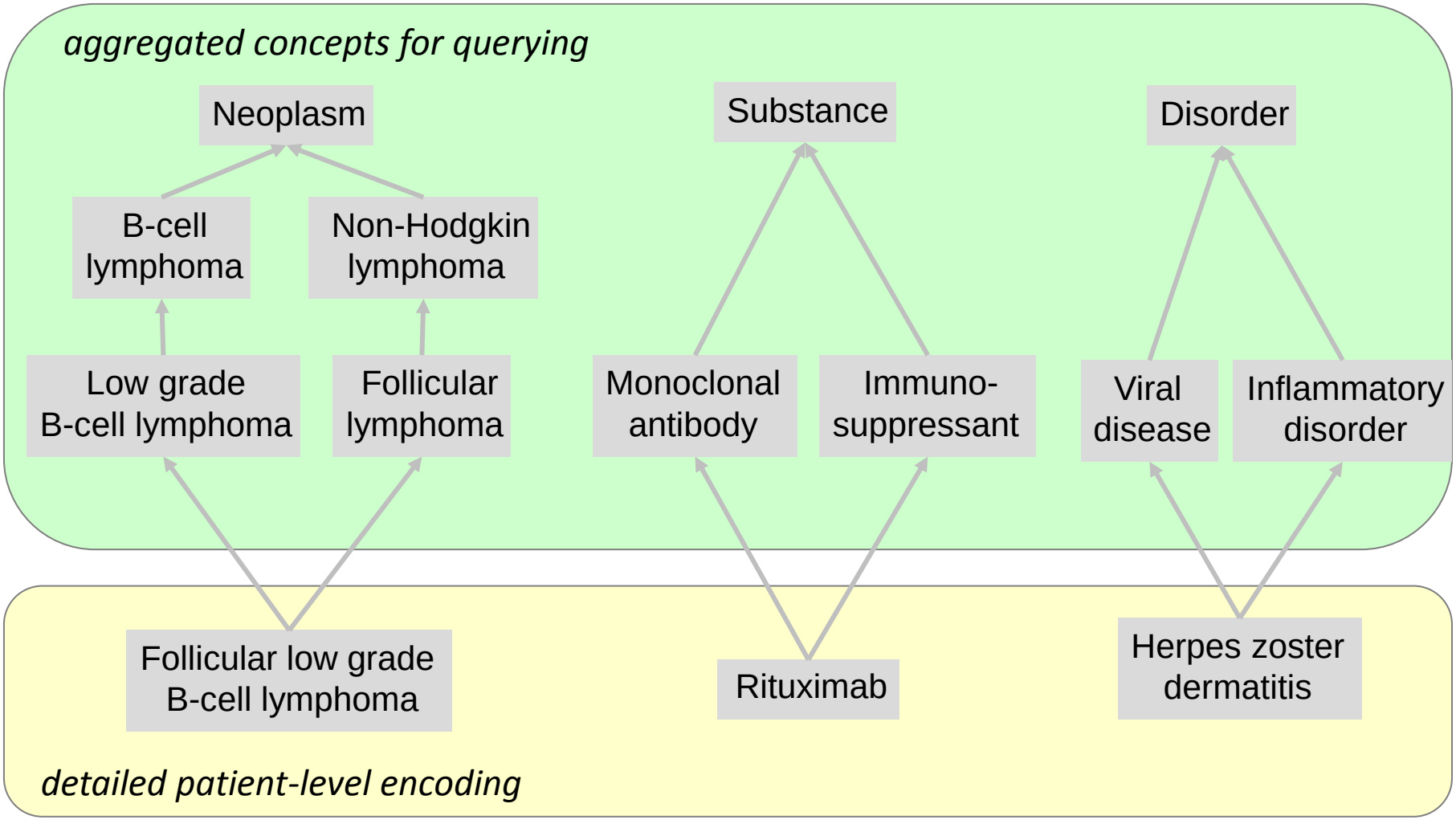
Interoperability ecosystem



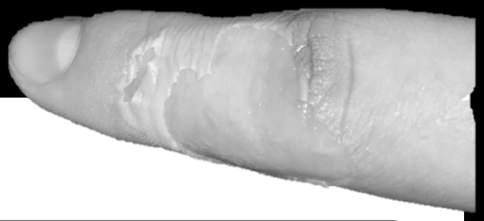
SNOMED CT: reference terminology



SNOMED CT – Structural benefits (I): Polyhierachies



SNOMED CT – Structural benefits (II): Co-ordination



Pre-coordination

"Verbrennung 2. Grades eines einzelnen Fingers"

```
211908006 |Deep partial thickness burn of a single finger (disorder)|
                                    ≡
<<< 29673001 |Second degree burn of single finger, not thumb (disorder)| :
{ 116676008 |Associated morphology| = 262588000 |Deep partial thickness burn
(morphologic abnormality)|,363698007 |Finding site| = 56213003 |Skin of
finger (body structure)| }
```

Post-coordination

"Verbrennung 2. Grades der Rückseite des rechten Zeigefingers"

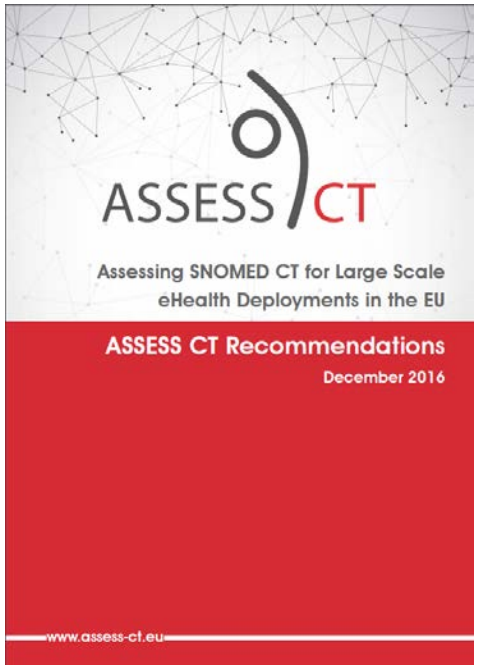
```
<<< 29673001 |Second degree burn of single finger, not thumb (disorder)| :
{ 116676008 |Associated morphology| = 262588000 |Deep partial thickness burn
(morphologic abnormality)|,363698007 |Finding site| = 37314006 | Skin
structure of dorsal surface of index finger (body structure) |, 272741003
|Laterality| = 24028007 |Right (qualifier value)| }
```

SNOMED CT – Challenges for MI-I

- Three major scenarios of use
  - SNOMED CT codes linked to (small) value sets for structured data entry
  - SNOMED CT as vocabulary for manual annotation of unstructured text
    - broad coverage of clinical jargon in medical texts
    - dealing with ambiguous expressions
    - known low inter-annotator agreement with large vocabularies (ASSESS-CT)
  - SNOMED CT as a vocabulary for machine annotations of clinical text

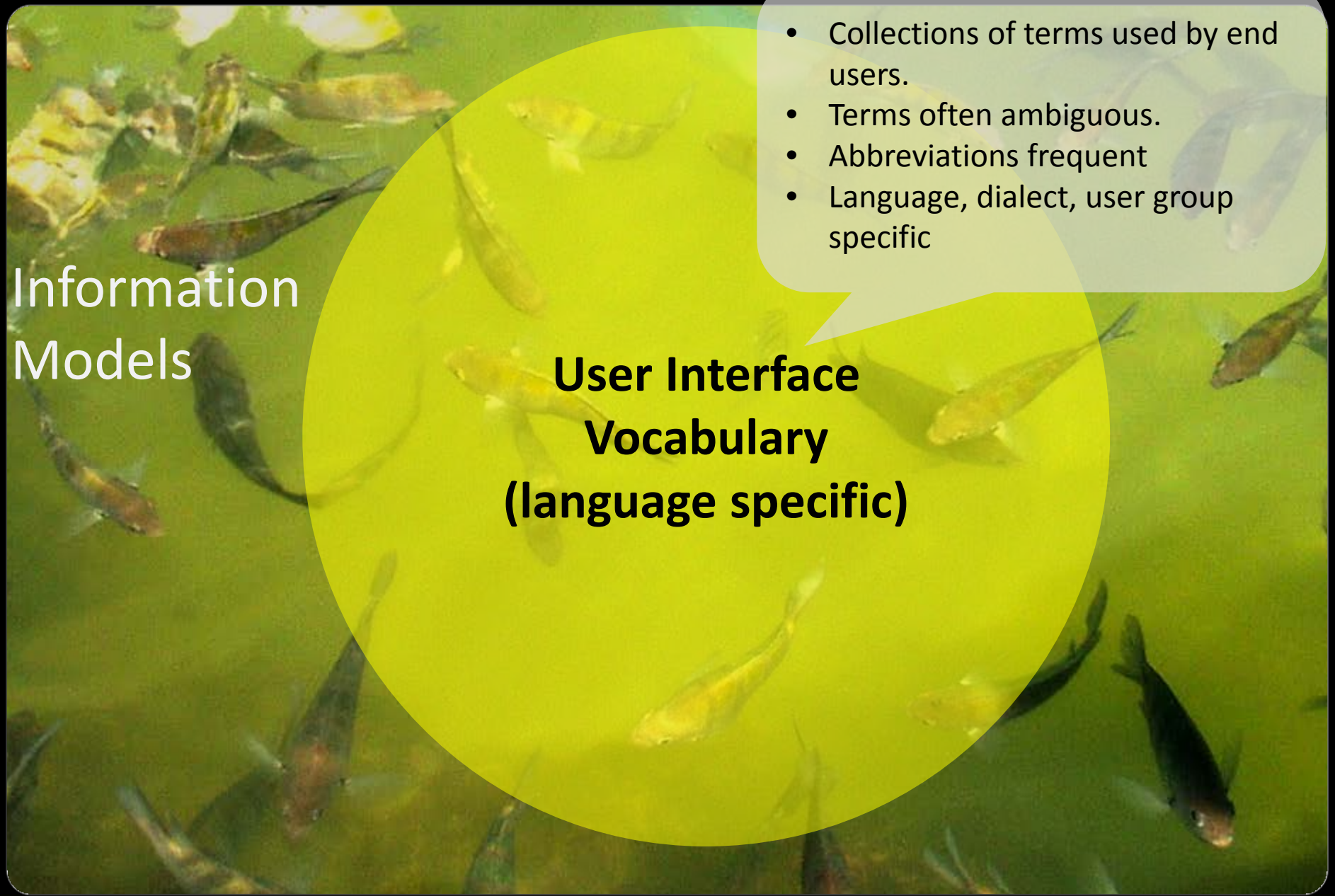
ASSESS CT recommendations

- "SNOMED CT is the best candidate for a core reference terminology for cross-border, national and regional eHealth deployments in Europe."
- SNOMED CT adoption must be part of a wider, strategy for optimising the benefits of semantic interoperability in health data
- The adoption of SNOMED CT should be realised incrementally rather than all at once
- Mechanisms should be established to facilitate and co-ordinate European Member State co-operation on terminology and semantic interoperability
- SNOMED CT should be part of a terminology ecosystem, including international aggregation terminologies and as wells as local user interface terminologies, which address multilingualism in Europe and clinical communication with multidisciplinary professional and lay language

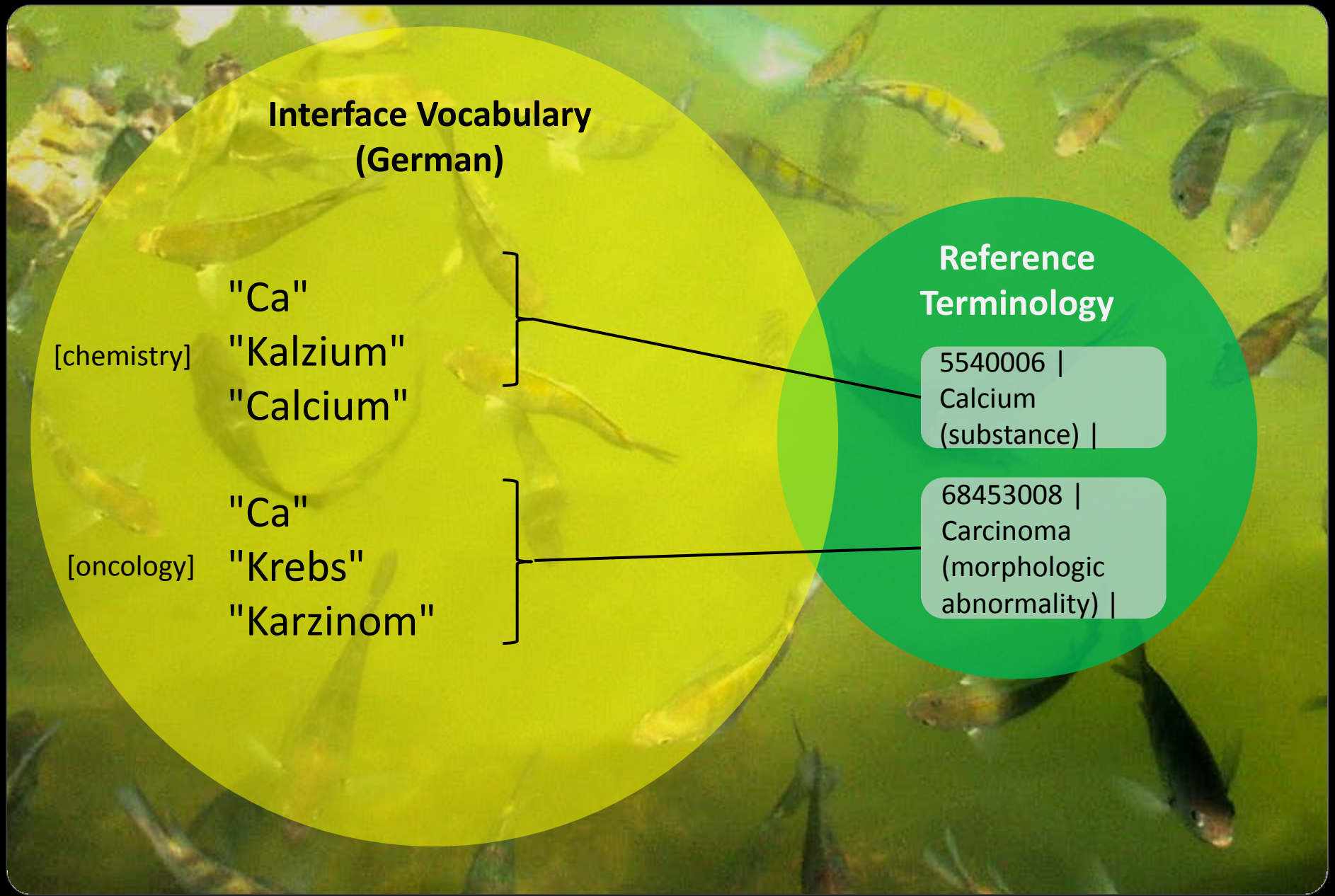


http://assess-ct.eu/fileadmin/assess_ct/final_brochure/assessct_final_brochure.pdf

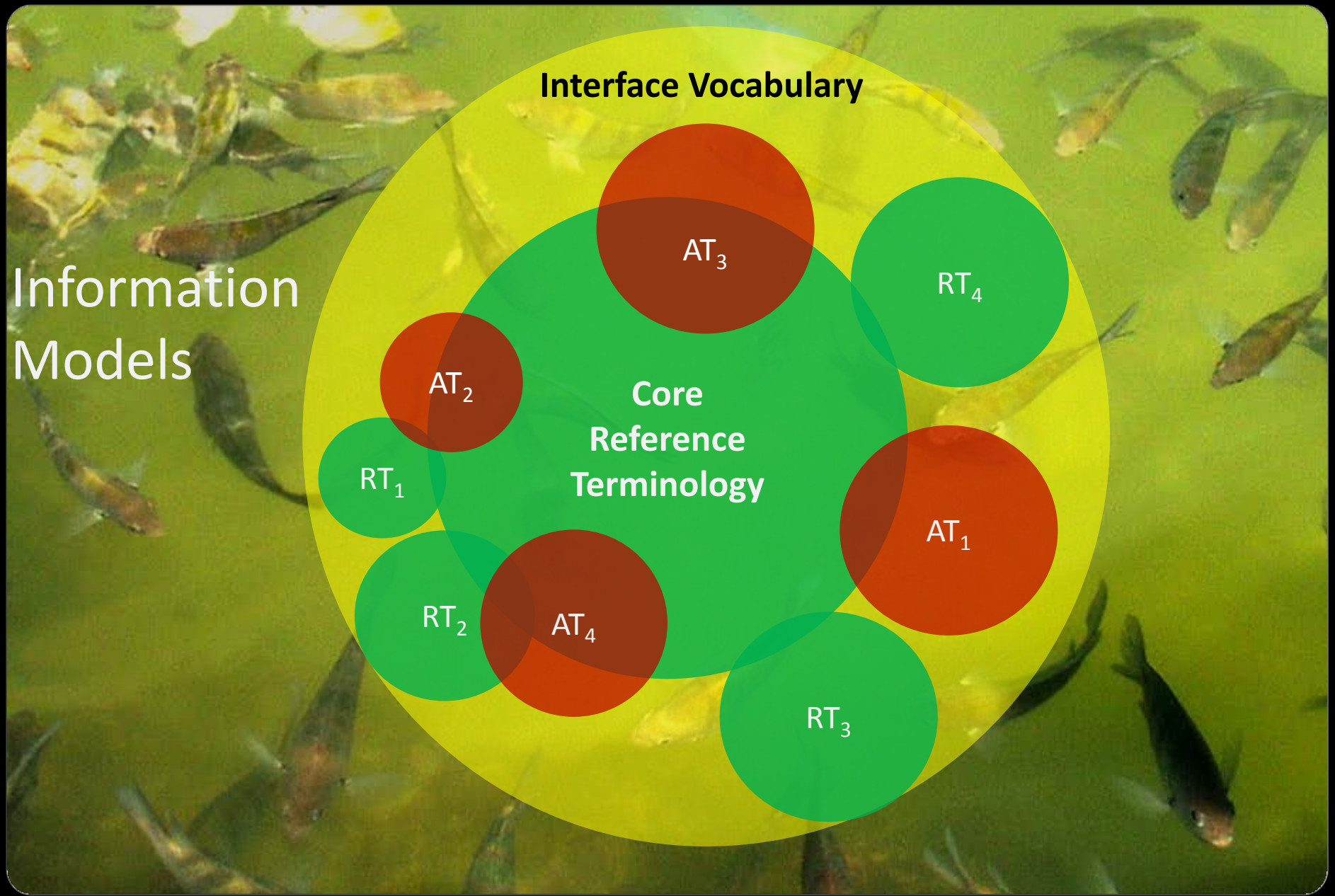
Interoperability ecosystem



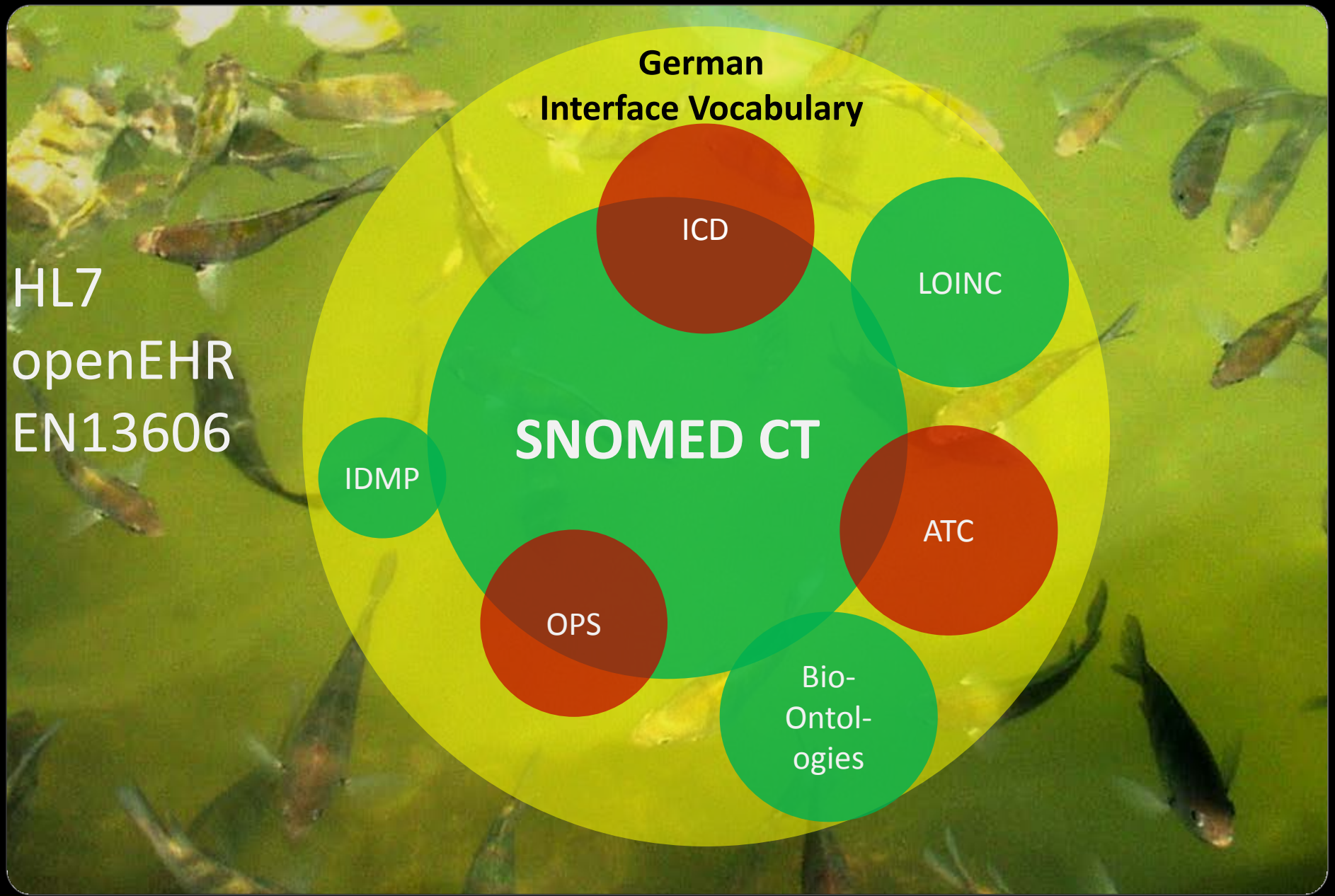
Interoperability ecosystem



Interoperability ecosystem



Interoperability ecosystem



SNOMED CT – Requirements for MI-I

- Acquire a SNOMED CT licence
  - Project-specific free research licence
  - Membership of SNOMED International (IHTSDO)
- Build and maintain German-language interface vocabulary
  - Joint effort of MI-I, in cooperation with international partners (Switzerland, Austria)
  - Use case driven
  - Crowdsourcing approach
  - In concert with advanced human language technology
- Experience: SNOMED CT Interface Vocabulary currently created at Medical University of Graz (Austria)

Preferred terms ≠ Interface terms

- Term occurrences in corpus of 30k cardiology discharge summaries (Med. Univ. Graz)

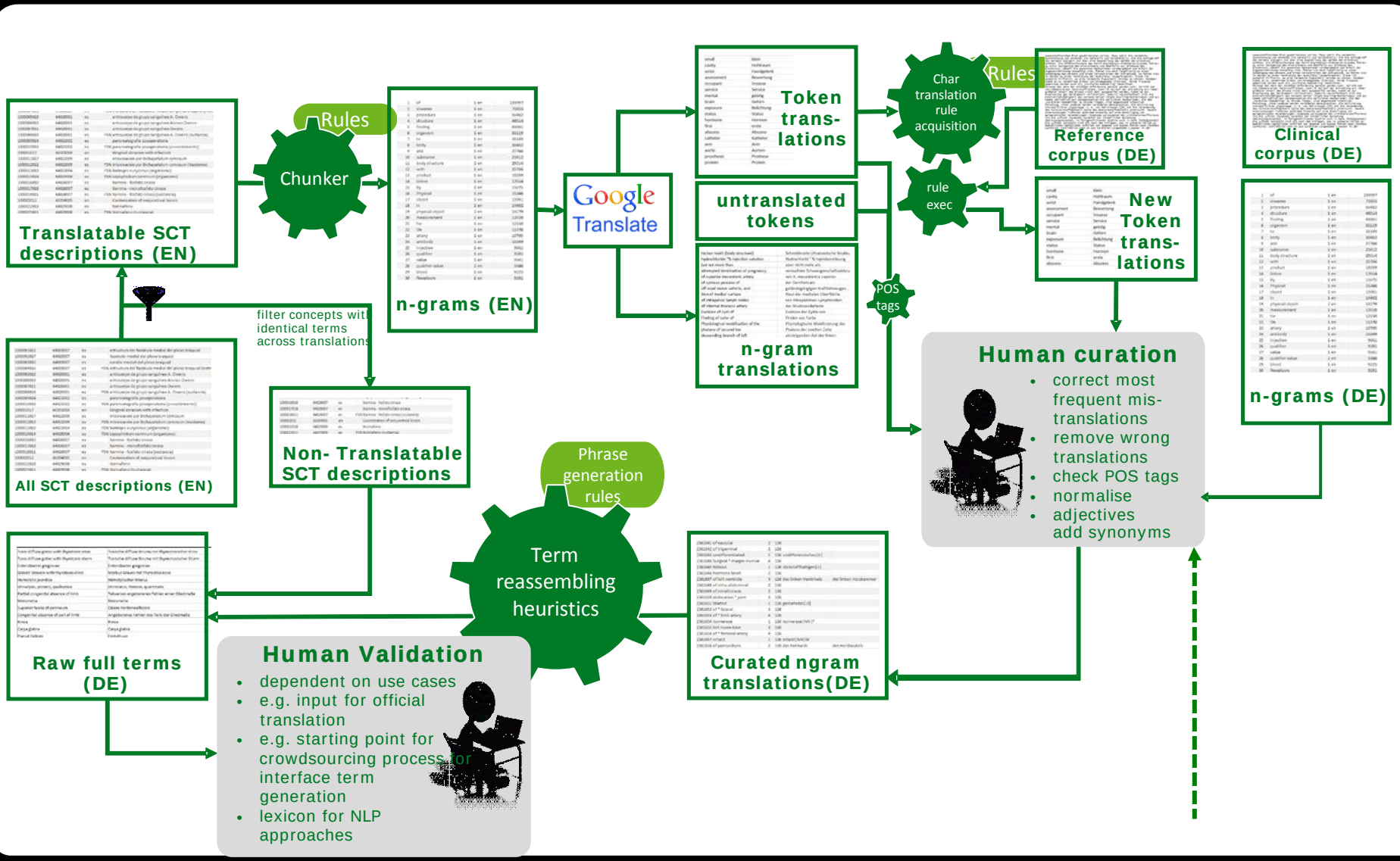
Preferred term (ICD, OPS)	Count	Interface term	Count
Aortenklappenstenose	3749	Aortenstenose	3126
Hirninfarkt	7	Schlaganfall	65
Elektrokardiogramm	0	EKG	12208
Koronare Herzerkrankung	331	KHK	18455
Nicht-ST-Hebungsinfarkt	498	NSTEMI	3839
Magnetresonanztomographie	2	NMR	17

MUG-GIT (Medical University of Graz – German Interface Vocabulary for SNOMED CT)

- Motivation: semantic EHR extracts for biomarker research (CBmed IICCAB*)
- Approach
  - Split of SNOMED CT terms into word n-grams
  - Manual editing of derived core vocabulary
  - Priority: term frequency and use cases
  - German terms from clinical corpora and terminologies
- Current state
  - Approx. 2 Million machine-generated interface terms, out of 95,000 n-gram translations
  - Core vocabulary maintained by medical students, guideline-driven (e.g. inclusion of word stems, spelling rules, acronyms only in context)

*Schulz S. Innovative Nutzung von Informationen für klinische Versorgung und Biomarkerforschung. http://goo.gl/wHMedz

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



Schulz S. Building an experimental German user interface terminology linked to SNOMED CT. Proc. SNOMED Expo 2017, Bratislava

Word 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

Suggestion: collaborative development of SNOMED CT interface dictionary within MI-I

- Rationale
  - Project on data integration / secondary use / semantic search: MI-I, Austria: ELGA, Switzerland: SNOMED CT introduction
  - Need for local vocabulary / value sets
  - Natural language processing of clinical texts
- Dictionary Crowdsourcing platform
  - Bottom-up development of German language interface terms linked to SNOMED CT
  - Term translations, context of use, example annotations
  - Recommender services, user rating
  - Joint ownership of content

Take home messages for MI-I projects

- **MI-I** as a **catalyst for SNOMED CT adoption** in DE
- Consider **SNOMED CT as a central part of a semantic ecosystem**, populated by other interfacing standards, e.g. IDMP, LOINC, ICD, HL7 etc.
- **Build a terminology task** force across all MI-I project and collaborate with other SNOMED CT users / adopters in the D-A-CH countries
- Focus on the **collaborative building of a German clinical interface vocabulary** linked to SNOMED CT, using state-of-the are crowdsourcing technology

MIRACUM Kick-off Symposium
22nd - 23rd February 2018
Erlangen

miracum

Medical Informatics for Research and Care in University Medicine

Stefan Schulz

Medical University of Graz (Austria)

purl.org/steschu

Acknowledgements:

semcare

ASSESS CT

SNOMED International

KAGes

SAP

FFG