



# From narratives to knowledge graphs. Principles for standards-based annotations of clinical narratives

UK Symposium on Knowledge Discovery and Data Mining  
June 14<sup>th</sup>, 2023



**Stefan Schulz**

- Medical University of Graz, Austria
- Averbis GmbH, Freiburg, Germany



# Background – Language and Semantics in Healthcare

# Biomedicine – area best covered by **interoperability** resources

## ■ Terminologies (not based on logic)

- Taxonomies, classifications, catalogues
  - International Classification of diseases (ICD-9, **ICD-10**, ICD-11)
  - ATC (Anatomical Therapeutic Classification System)
  - NCBI Taxonomy (biological)
- Thesauri
  - MeSH – Literature indexing
  - MedDRA – Drug Regulation
  - NCIthesaurus (cancer docu)
  - RxNorm (drugs)
  - **LOINC** (lab and other obser

## ■ Data / Information models

- **HL7 FHIR**
- EN 13606
- openEHR

## ■ Ontologies (based on logic)

- Ontology-based terminology
  - **SNOMED CT** (electronic health records)

- Identifiers
- Labels
- Synonyms
- Translations
- Definitions
- Axioms
- Documentation
- Templates

- ologies
- ne Ontology (activities, processes, sites)
- quence Ontology (nucleotides, protein sequences)
- EBi (chemical entities)
- O (human phenotypes)
- IA (anatomy)
- ses”**
- ogues of similar entities with instance s by ontologies)
- UNIPROT (Proteins)
- Reactome (biological pathways)
- BRENDA (enzymes)

# Health care:

## Purpose of biomedical interoperability resources / standards

- Healthcare and medicine:
  - **Routine coding, e.g. for reimbursement or controlling** (diagnoses, procedures)
    - Major source of bias: selective, coarse-grained, unprincipled
  - **Mortality and Morbidity statistics** (e.g. ICD for diseases at WHO level)
  - **Clinical registries** (e.g. tumour documentation)
  - **Drug regulatory activities**
  - Clinical decision support
  - Clinical research
  - Standardisation of clinical data in electronic health records (EHRs)



**Clinical coding / data curation  
is expensive!**

# Most clinical information is contained in clinical narratives

... in the local natural language



Porto Alegre  
Brazil



Graz  
Austria

Paciente G1PO, IG de 38 sem 4 dia(s), TS A+, interna por bolsa roita há mais de 18hs, recebendo penicilina. Evolui para Parto Eutócico com episiotomia em 27/06/2007 22:24 hs. Nasce RN APGAR 10/10, MASC, 3060 G. Exames: Toxo IGG e IGM neg VDRL neg EQU neg UROC: ausência de crescimento bacteriano. Hemograma 198mil plaq; Hb 13,1; LT 12,5 (75% seg) Em condições de alta, amamentando, útero contraído, lóquios fisiológico, sinais vitais estáveis, FO com bom aspecto. Recebe as orientações abaixo. ORIENTAÇÕES NA ALTA: # AMAMENTAÇÃO EXCLUSIVA POR 6 MESES; # TOMAR AS MEDICAÇÕES PRESCRITAS (SULFATO FERROSO 300MG 3X/DIA POR 90 DIAS, LONGE DAS REFISÇÕES, COM SUCO DE LARANJA; PARACETAMOL 750 MG 6/6HS SE DOR); # ORIENTO ANTICONCEPÇÃO; # RETORNAR À EMERGÊNCIA DESTE HOSPITAL SE FEBRE, SANGRAMENTO AUMENTADO OU OUTRAS INTERCORRÊNCIAS. # NÃO É NECESSÁRIO RETIRAR OS PONTOS. # LAVAR FO 3X/DIA COM ÁGUA E SABÃO DE GLICERINA.

\* Anamnese und klinische Symptomatik  
Stat. Übernahme vom LKH Fürstenfeld wegen neuerlicher Dyspnoe bei bek. dil. CMP u hochgr. MINS zur CA und Mitraclip /erztransplant Evaluierung. Bei dem Patienten besteht der St.p. 2x Simdax Therapie im Okt 2013.  
\* Physikalischer Status  
48 jähr.Patient, deutl. reduz. AZ, normaler EZ. Cor: Ht rh, nc, Systolikum mit p.max. über dem Erbschen Punkt mit Fortleitung in die Axila  
Pulmo: VA bds., feuchte RGs re>li  
Abdomen: BD weich, kein DS  
Extremitäten: ausgeprägte Knöchelödeme bds.  
Herr DI Max Mustermann wurde aufgrund einer neuerlichen Dyspnoesymptomatik bei bek. dilat. CMP und hochgrad. MINS zur weiteren Evaluierung stat. vom LKH Fürstenfeld übernommen.

# Clinical language: compact, sloppy, contextualised

Phenomenon	Example	Elucidation
Telegram style	“left PICA stroke, presented to ED after fall”	Incomplete sentences, sketchy style
Colloquialisms	“pothole sign”, “snorkel”	Milieu-specific sub-languages
Ad-hoc abbreviations	“infiltr”	Truncation (“infiltrated mucosa”)
Ambiguous short forms	“RTA”	“Road traffic accident”, “Renal-tubular acidosis”
Short forms of regional or local scope	“LDS Hospital” “St. p.”	“Latter-Day-Saints Hospital” (and not “Leak Detection System”) “Status post” = “History of”
Conventionalized Latin abbreviations	“V mors can dig V dext”	“Vulnus morsum canis digiti quinti dextri” (in some European languages)
Numeric codes	“45, 46 with crowns”, “VI palsy”, “2-2-2”,	Tooth numbers, cranial nerves, dose frequencies
Spelling errors, typos	“Diabtes”, “Astra-Seneca”, “Hipotireose”,	accidental (quick typing) or systematic (e.g. 2 <sup>nd</sup> language speakers)
Spelling variants	“Esophagus”, “Oesophagus”	e.g. American vs. British English
Single noun compounds	“Ibuprofenintoxikation”	Non-lexicalized long words (in languages such as German, Swedish)
Anaphora	(i) “adenoCa rect pN+MX G2 (...). tumor excised in toto” (ii) “no blood in stomach (...). mult mucosal erosions ”	(i) “Tumor” coreferential to adenocarcinom described in left context (ii) “mucosal erosions” refined to “erosions of gastric mucosa”
Negations	“No evidence of pneumonia” “Pulmones: nihil”, “metastasenfrei”	non-standard, jargon-like
Epistemic contexts	“susp MI, DD lung embolism”	suspected diagnosis, differential diagnosis
Temporal contexts	“h/o Covid-19”, “Streptokokkenangina 06/16”	“history of” Coarse-grained references to dates (mm/yy)
Other contexts	(i) father: pancreas ca” (ii) “refrained from resuscitation”	(i) family history (ii) plans not executed

# Making health record data interoperable: narratives → ontology-based data

Physical examination on admission revealed purpura of the upper and lower extremities, **swelling of the gums and tonsils**, but no symptoms showing the complication of myasthenia gravis. Hematological tests revealed leucocytosis: WBC count 68 700/ $\mu$ l (blasts 11.5%, myelocytes 0.5%, bands 2.0%, segments 16.0%, monocytes 65.5%, lymphocytes 4.0%, atypical lymphocytes 0.5%), Hb 7.1 g/dl (reticulocytes 12%) and a platelet count of  $9.1 \times 10^4$ / $\mu$ l. A bone marrow aspiration revealed hypercellular bone marrow with a decreased number of erythroblasts and megakaryocytes and an increased number of monoblasts

At the level of fine-grained entity types  
(not broad categories like disease procedure, drug)

Parents

- Pharyngeal swelling (finding)
- Swelling of head (finding)
- Tonsil finding (finding)

Swelling of tonsil (finding)  
SCTID: 442394001  
442394001 | Swelling of tonsil (finding) |  
en Swelling of tonsil (finding)  
en Swelling of tonsil

Finding site → Palatine tonsillar structure  
Associated morphology → Swelling

Children (2)

- Swelling of left tonsil (finding)
- Swelling of right tonsil (finding)

Parents

- Acute digestive system disorder (disorder)
- Acute inflammatory disease (disorder)
- Gingivitis (disorder)

Acute gingivitis (disorder)  
SCTID: 31642005  
31642005 | Acute gingivitis (disorder) |  
en Acute gingivitis (disorder)  
en Acute gingivitis  
en Acute gingival inflammation

Finding site → Gingival structure  
Associated morphology → Acute inflammation  
Clinical course → Sudden onset AND/OR short duration

Children (4)

- Acute gingivitis due to non-plaque induced gingival disease (disorder)
- Acute pericoronitis (disorder)
- Acute ulcerative gingivitis (disorder)
- Vincent's laryngitis (disorder)

# Desideratum: narrative data → ontology-based data

The screenshot shows a web interface for text-to-ontology conversion. At the top, there are tabs for 'Text' and 'Documents'. Below that, there are dropdown menus for 'DETECT LANGUAGE' (with options ENGLISH, SPANISH, FRENCH) and a bidirectional arrow, followed by another set of dropdowns for 'ENGLISH', 'SPANISH', and 'SNOMED CT'. The main content area is split into two columns. The left column contains a clinical narrative, and the right column contains a list of SNOMED CT codes.

Physical examination on admission revealed purpura of the upper and lower extremities, swelling of the gums and tonsils, but no symptoms showing the complication of myasthenia gravis. Hematological tests revealed leucocytosis: WBC count 68 700/ $\mu$ l (blasts 11.5%, myelocytes 0.5%, bands 2.0%, segments 16.0%, monocytes 65.5%, lymphocytes 4.0%, atypical lymphocytes 0.5%), Hb 7.1 g/dl (reticulocytes 12%) and a platelet count of $9.1 \times 10^4$ / $\mu$ l. A bone marrow aspiration revealed hypercellular bone marrow with a decreased number of erythroblasts and megakaryocytes and an increased number of monoblasts	419620001 110714004 65124004 113279002 116223007 91637004 252275004 111583006 767002 [68700] 271040006 [11.5] 313696224 [0.5] 313696667 [2.0] 313696009 [16.0] 271037006 [65.5] 271036002 [4.0] 271036013 [0.5] 365809007 [7.1] 45995003 [12] 365632008 [91000] 49401003 76197007 14016003 420510009 103213002 53945006 35105006
--	--

But language is not linear...



# Knowledge graphs rooted in interoperability standards: preferred representation of clinical information

- Describing a clinical document
- Describing a clinical treatment episode
- Describing lifelong clinical history of a patient
- Describing many (anonymous) patients

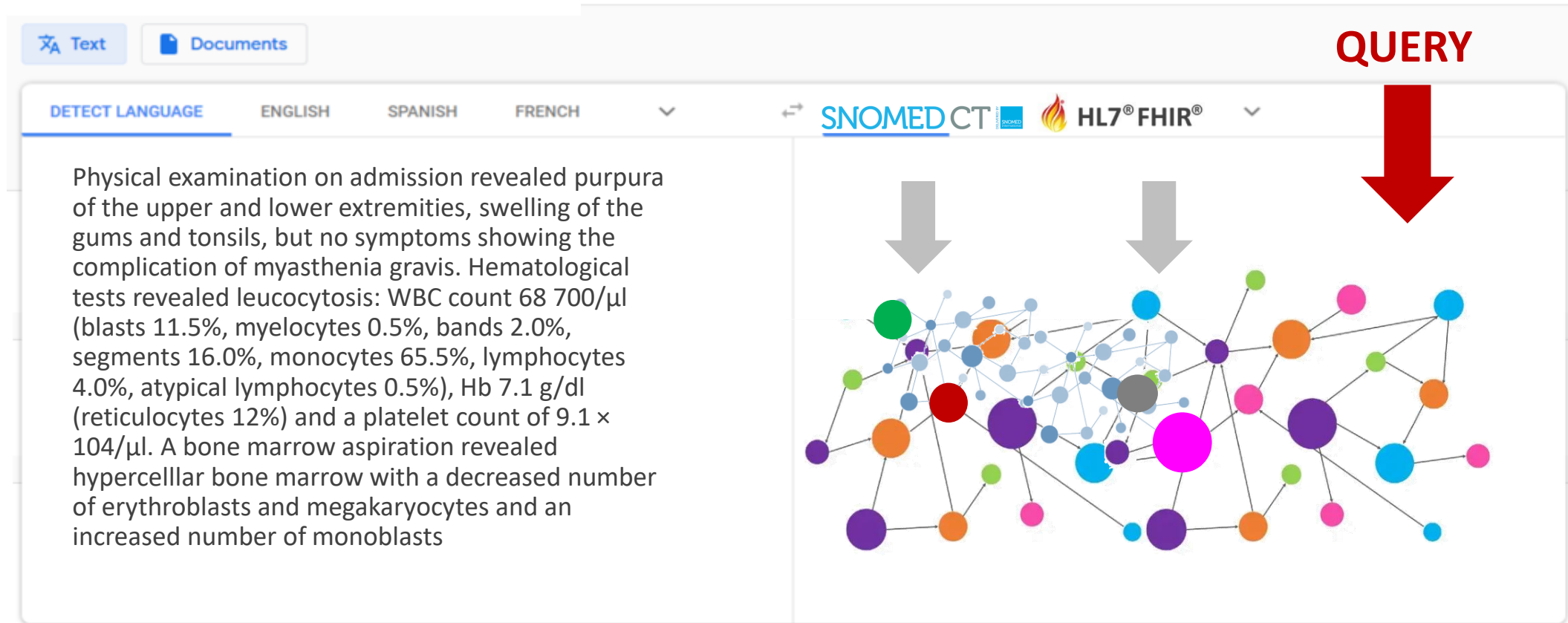
Text Documents

DETECT LANGUAGE ENGLISH SPANISH FRENCH

SNOMED CT HL7® FHIR®

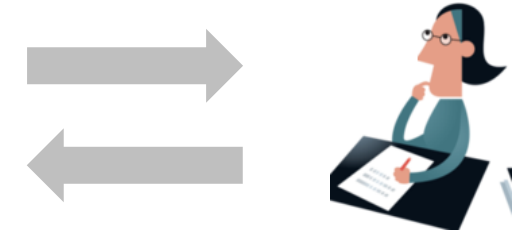
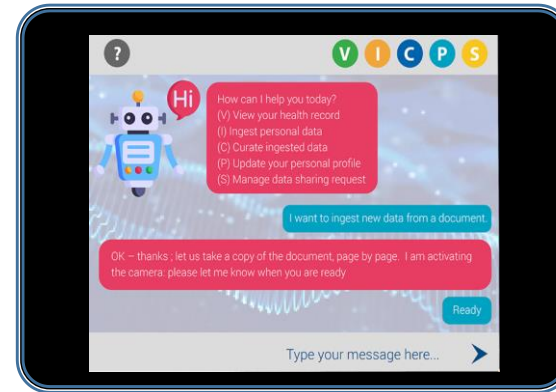
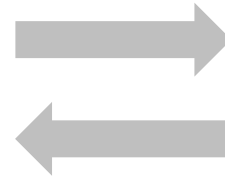
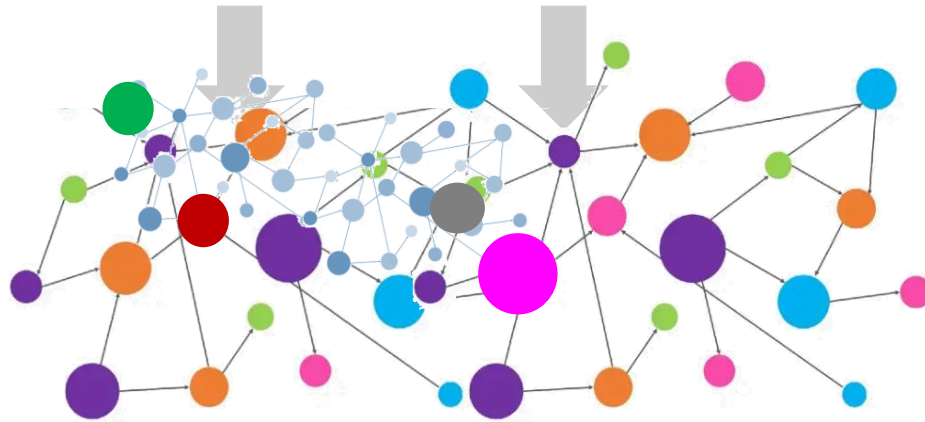
Physical examination on admission revealed purpura of the upper and lower extremities, swelling of the gums and tonsils, but no symptoms showing the complication of myasthenia gravis. Hematological tests revealed leucocytosis: WBC count 68 700/ $\mu$ l (blasts 11.5%, myelocytes 0.5%, bands 2.0%, segments 16.0%, monocytes 65.5%, lymphocytes 4.0%, atypical lymphocytes 0.5%), Hb 7.1 g/dl (reticulocytes 12%) and a platelet count of  $9.1 \times 10^4$ / $\mu$ l. A bone marrow aspiration revealed hypercellular bone marrow with a decreased number of erythroblasts and megakaryocytes and an increased number of monoblasts

QUERY



# Scenarios of use of clinical knowledge graphs

SNOMED CT   HL7® FHIR®



Health data curators

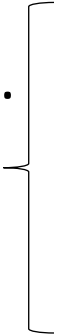
- Primary use:

- Retrieval
- Summarization
- Information extraction
- Input for decision support
- Data visualization, navigation

- Secondary use:

- Retrospective research
- Patient recruiting
- Training of predictive models
- Training of decision support systems
- Quality improvement

# Methods and resources to build knowledge graphs from clinical narratives

- Natural language processing (NER + concept extraction, relation extraction, coreference resolution, text classification,... )
- Methods
  - Rules
  - Lexicon lookup
  - Machine learning
  - Deep learning
  - Large language models
- Resources
  - Terminologies, ontologies etc.
  - **Annotated text corpora**
  - Enhancing lexicons
  - Training ML models
  - Training / Domain fine-tuning DL models
  - Exploiting large language models
  - Benchmarking

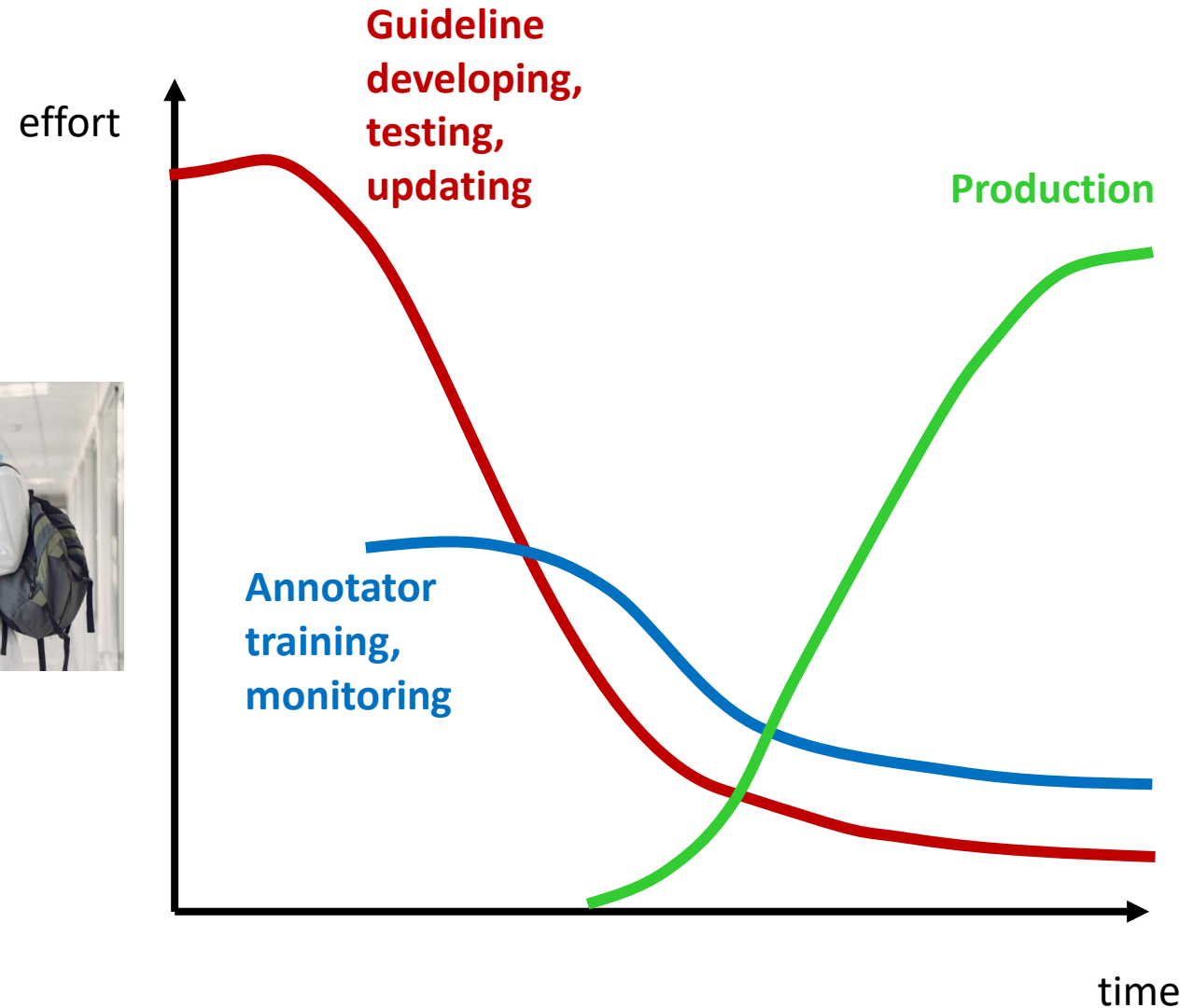
# Principles of Clinical Document annotation

# Known issues of clinical corpus annotation

- Domain expertise
- Extensive training
- Inter-annotator variability
- Annotation fatigue
- Ambiguities
- Time constraints
- Success factors:
  - Rigorous annotation guidelines
  - Good tooling
  - Repeated training sessions
  - Adjudication between annotators
  - Quality checks (inter-annotator agreement)
  - communication channels



medical students



# Annotation guideline and Annotation tool



## Annotation guideline for semantic annotations of clinical narratives based on SNOMED CT and FHIR

Stefan Schulz<sup>1,2</sup>, Akhila Naz Kuppasery<sup>1</sup>, Alexander Beger<sup>1</sup>, Sareh Aghaei<sup>1</sup>, Daniel Dür<sup>1</sup>, Larissa Hammer<sup>1</sup>, Kristian Kankainen<sup>3</sup>, Markus Kreuzthaler<sup>1</sup>, NN, NN...

<sup>1</sup>Medical University of Graz, Austria

<sup>2</sup>Averbis GmbH, Freiburg, Germany

<sup>3</sup>Tallin University of Technology, Estonia

Version 2023-06-14

Guests are welcome! Feel free to drop comments.  
E-mail contact: stefan.schulz@medunigraz.at

In progress: consolidation of practical content for insertion into AIDAVA D4.2

1. Introduction	1
2. Background	2
2.1. Annotation strategies	2
2.2. Related work	4
3. Objectives	5
4. Tools and resources	6
4.1. Naming and graphical conventions	6
4.2. Annotation vocabularies	6
4.2.1. SNOMED CT	6
4.2.2. HL-7 FHIR	6
4.2.3. LOINC	7
4.3. INCEPTION	7
5. Basic assumptions and decisions	7
5.1. Conceptual model	7
5.2. Principles	7
5.3. Preferences regarding SNOMED content	10
5.3.1. Core (focus) concepts	1
5.3.2. Modifying concepts	12
5.3.3. Product concepts	12
5.3.4. Concepts from HL7 value sets	12
5.3.5. Combination of annotations	13
5.4. Concepts that require values	14
5.5. Relations, predicates, and operators	16
5.5.1. Negation, Uncertainty, clinical status, and severity	16

<https://bit.ly/3X4McGC>

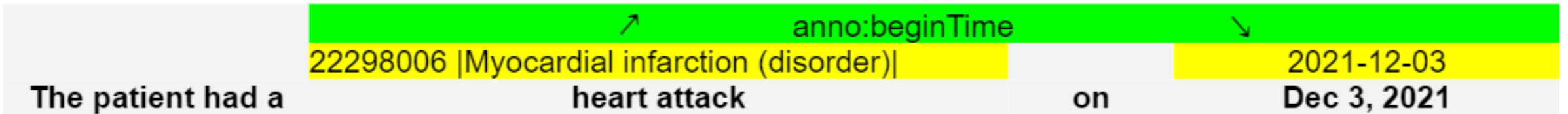
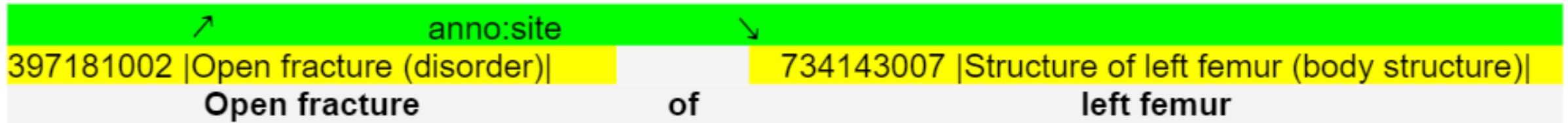
The screenshot displays the INCEPTION web application interface. The main window shows a clinical narrative with various parts annotated with colored boxes and lines representing semantic relationships. The narrative includes text like "DEKURS DER TUMORERKRANKUNG" and "Operation(en), Histologie(n), Immunhistologie(n), Molekulare(s) Profil(e):". Annotations include "Surgical procedure", "Histologic test", "Molecular genetic test", "Partial mastectomy", "Excision of axillary lymph node", "Excision", "Procedure status", "American Joint Committee on Cancer pT2", "American Joint Committee on Cancer pN1a", "American Joint Committee on Cancer grade G2", "American Joint Committee on Cancer R0", "Clinical stage finding", "Proliferation marker protein Ki-67", "Histologic test", "Solid ductal carcinoma in situ of breast", "American Joint Committee on Cancer pTis", "Immunohisto", "Immune (qualifier value)", "High", "High", "High", "Histologic test", "Oestrogen receptor positive tumour", "Progesterone receptor positive tumour", and "HER2-positive carcinoma of breast". The right sidebar shows the "Annotation" panel with fields for "Layer" (Custom MCN), "Text" (axill. Dissektion), "Comment", "Concept" (234262008 | Excision of axillar), and "Short" (Excision of axillary lymph nod).

# Basic annotation principles

- Two level annotation – text spans and relations
- Deep annotation: maximum granularity
- Flexible spans (subword to multiword): determined by vocabulary
- Longest match preference: benefit from pre-coordinated concepts
- Close-to-text: no interpretation
- Co-reference annotations
- Standards- and ontology-aware annotation

# Two level annotation

1. Text spans, annotated with codes or literals 
2. Linkage of text spans by binary predicates 





# Deep annotation

Annotations exploit the whole depth of the annotation vocabulary

No “entity-type” annotation

12236201000119103 |Conjunctivitis of right eye (disorder)|

**Unilateral conjunctivitis right**

Not:

Disorder

**Unilateral conjunctivitis right**

# Flexible annotation spans / longest match principle

- Annotations spans determined by annotation vocabulary
- Preference given to longest match (precoordinated concepts)
- No determination of spans by NER before annotation

28576007 |Open fracture of femur (disorder)|

The femur exhibited an open fracture



anno:site



76505004 |Thumb structure (body structure)|

The thumb

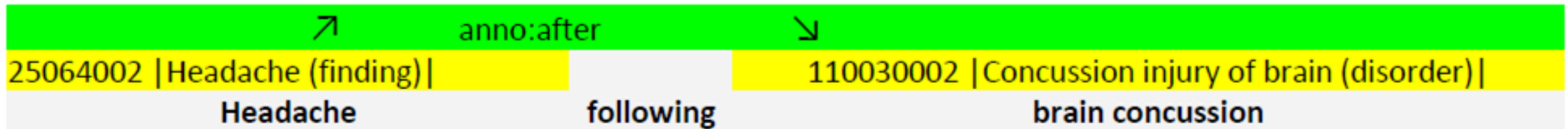
had an

397181002 |Open fracture (disorder)|

exposed fracture

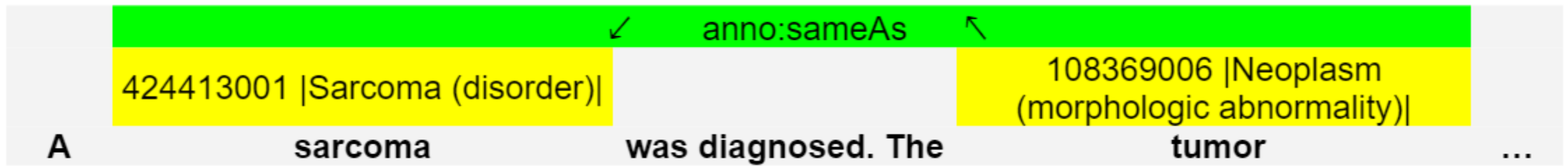
# Close-to-text: no interpretation of content

- Only annotate what is explicitly stated, not what might be medically plausible

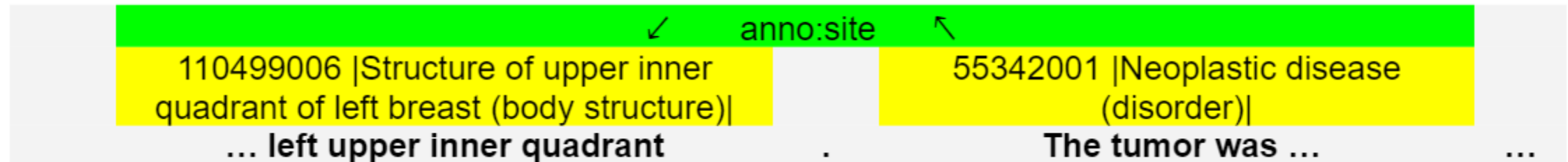


# Coreference annotations

- Nominal / pronominal anaphora via same as







- Bridging anaphora via predicates



# Standards- and ontology-aware annotations

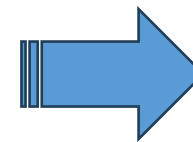
- Annotation vocabulary is related to domain standards (SNOMED CT, FHIR)
- Set of annotation predicates defined by the guideline
- Mapped to relations or relation chains in underlying standards

Predicate	Domain constraints	Predicate semantics	Range constraints
inFamily	<<404684003  Clinical finding (finding)	INV(FamilyMemberHistory.condition)    FamilyMemberHistory.relationship 	<<303071001  Person in the family
		INV(246090004  Associated finding (attribute) )    408732007  Subject relationship context (attribute)  	
verificationStatus	<<404684003  Clinical finding (finding)  OR <<272379006  Event (event)	INV(Condition.code)    Condition.verificationStatus 	410590009  Known possible 415684004  Suspected 410592001 Probably present 410605003  Confirmed present 410594000  Definitely NOT present 410516002  Known absent 723510000  Entered in error 261665006  Unknown
		INV(246090004  Associated finding (attribute) )    408729009  Finding context (attribute)  	

# Document preprocessing before annotation (I)

- Using existing concept / relation tagger for pre-annotation:

The screenshot displays the HEALTH discovery interface. At the top, there are tabs for 'Texteingabe' and 'Text Analyse Ergebnisse'. Below this is a navigation bar with various categories: Anatomy, ClinicalSection, ClinicalSectionKeyword, Concept (checked), Date, Diagnosis, DocumentAnnotation, EstrogenReceptor, HER2, Morphology, PatientInformation, ProgesteroneReceptor, TNMGrading (checked), TNMMetastasis (checked), TNMNode (checked), and TNMTumor (checked). The main content area shows a document titled 'DEKURS DER TUMORERKRANKUNG' with several paragraphs of text. A search bar and a filter dropdown are visible on the right side of the document. The filter is set to 'TNMGrading' and shows results for 'G2' with a begin position of 233, an end position of 235, and a value of 'G2'. Below this, the 'Concept' filter is also set to 'G2', showing a begin position of 233, an end position of 235, a conceptID of 1228850007, a dictCanon of G2, a matchedTerm of G2, a source of SNOMED\_CT\_GIT, and a uniqueID of SNOMED\_CT\_GIT:1228850007.



# Document preprocessing before annotation (II)

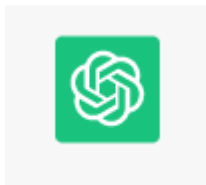
- Text cleansing by generative language models (experimental):

## Original text:

"Z.n. TE eines exulc. sek. knot. MM Pes sin. dors. Grad IV 2,4 mm Tumordurchm. Sentinell LK ing. li. tumorfr. (ED Mai 2017 )"



## Prompt:



Translate this diagnosis text into a text that general practitioners can understand without errors and short forms:



„Der Patient hatte eine Operation, bei der ein geschwürartiger sekundärer Knoten eines malignen Melanoms (MM) am linken Fußrücken, auf der Rückseite, Stadium IV, mit einem Tumordurchmesser von 2,4 mm entfernt wurde. Der Wächterlymphknoten (Sentinell LK) auf der linken Seite wurde untersucht und es wurden keine Tumorzellen gefunden (Untersuchungsdatum: Mai 2017)“

Training of AI models with annotated corpora



Clinical knowledge graphs as target structure

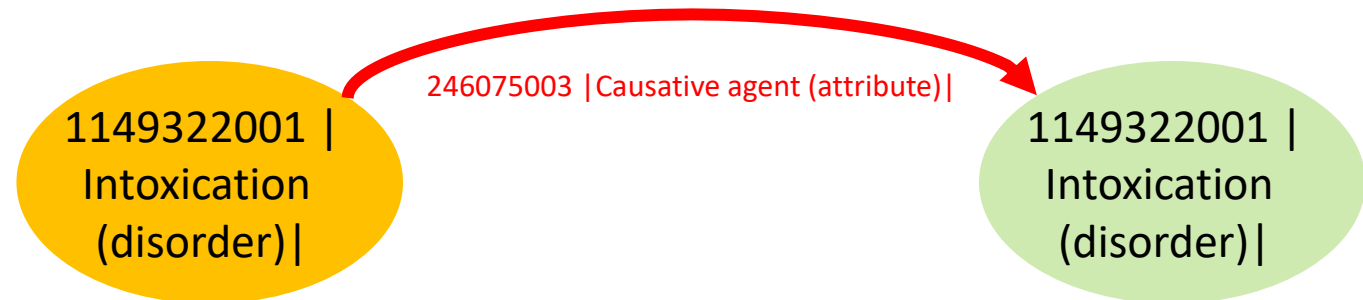
# Knowledge graph postprocessing after text analysis

- Exploitation of axiomatic structure of annotation ontology (e.g. SNOMED CT)

“Amphetamin intoxication”

```
EquivalentClasses(  
  :45421006 | Amphetamine intoxication (disorder)|  
  ObjectIntersectionOf(  
    :1149322001 | Intoxication (disorder)|  
    ObjectSomeValuesFrom(  
      :609096000 | Role group (attribute)|  
      ObjectSomeValuesFrom(  
        :246075003 | Causative agent (attribute)|  
        :703842006 | Amfetamine (substance)| ) ) ) ) )
```

“the intoxication  
was caused by  
Amphetamin”



# Identification of semantic equivalences

Condition

- verificationStatus
- code 390926006 | Suspected gallstones (situation) |
- bodySite

Condition

- verificationStatus unconfirmed
- code 235919008 | Gallbladder calculus (disorder) |
- bodySite

Condition

- verificationStatus unconfirmed
- code 313413008 | Calculus finding (finding) |
- bodySite 3578005 | Structure of body of gallbladder (body structure) |

Condition

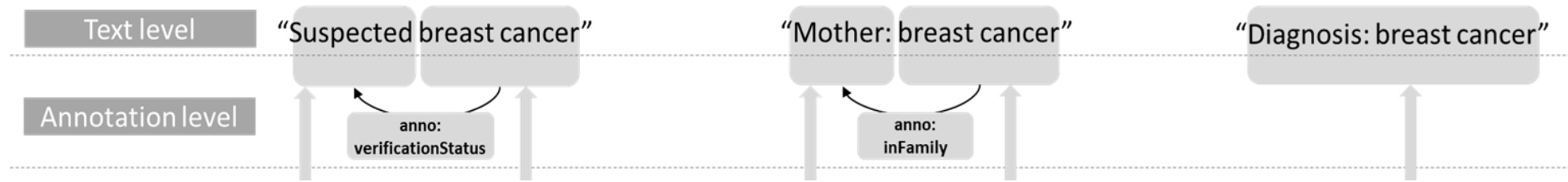
- verificationStatus unconfirmed
- code 56381008 | Calculus (morphologic abnormality) |
- bodySite 3578005 | Structure of body of gallbladder (body structure) |

Condition

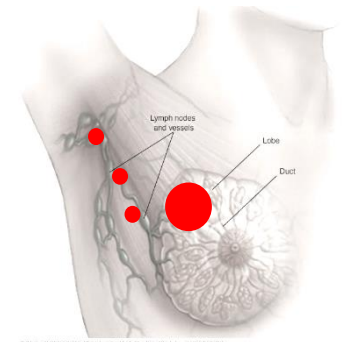
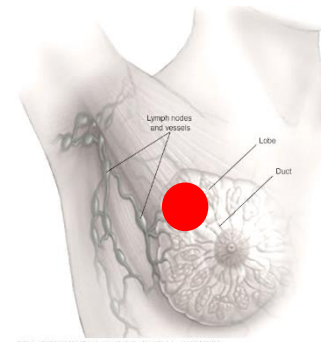
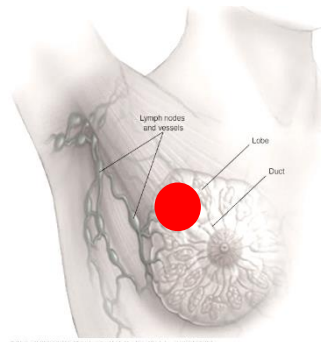
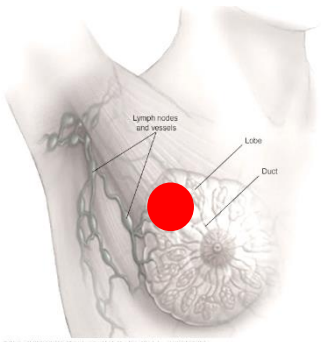
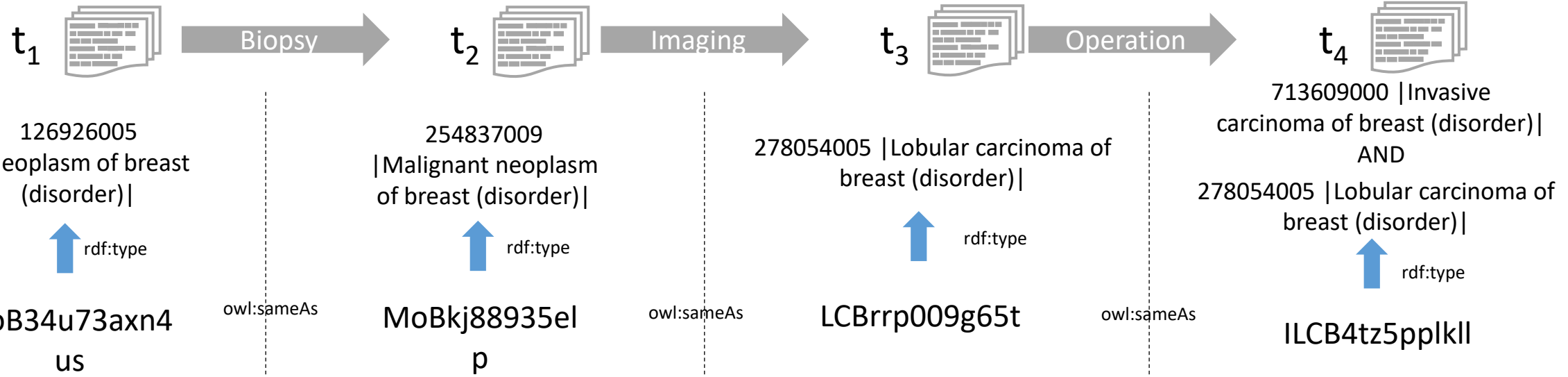
- verificationStatus
- code
- bodySite

```
EquivalentClasses(  
  :41769001 |Disease suspected (situation)|  
  ObjectIntersectionOf(  
    :444433005 |Suspected clinical finding (situation)|  
    ObjectSomeValuesFrom(  
      :609096000 |Role group (attribute)|  
      ObjectIntersectionOf(ObjectSomeValuesFrom(  
        :246090004 |Associated finding (attribute) | :64572001 |Disease (disorder)|  
        ObjectSomeValuesFrom(:408729009 |Finding context (attribute) | :415684004 |Suspected (qualifier value)|  
        ObjectSomeValuesFrom(:408731000 |Temporal context (attribute) | :410512000 |Current or specified time (qualifier value)|  
        ObjectSomeValuesFrom(:408732007 |Subject relationship context (attribute) | :410604004 |Subject of record (person) | ))))
```

# Open issue: OWL individuals vs. OWL classes in Knowledge Graph





# Open issue: Identity management



Lobular carcinoma of breast

Invasive Lobular carcinoma of breast

# Summary / Outlook

- Much clinical information only in narratives → NLP mandatory
- Biomedical NLP can highly benefit from existing semantic resources
  1. Ontologies (description of entity types): Definitions / Axioms
  2. Terminologies (description of natural language): labels, synonyms
  3. Information Models (Instance-level templates, link to 1. and 2.)
- Annotated corpora are essential for training and benchmarking NLP  
- Proposal of “deep” annotation principles for clinical narratives. Resulting annotation knowledge graph used for NLP / DL system training and benchmarking
- New perspectives regarding pre-annotation and KG post-processing, e.g.
  - Generative language models
  - Identification of semantically identical expressions
  - Identification of identical instances
- Application and assessment of annotation principles:
  - EU AIDAVA – Narratives → KG (Dutch, Estonian, German text)
  - GeMTeX – German Medical Text Corpus
- Cooperations
  - Technical University of Munich (Germany), University of Erlangen - Nürnberg (Germany), German Research Centre for Artificial Intelligence, University of Manchester (UK), University of Murcia (Spain), University of Chiang Mai (Thailand)

SNOMED CT 

 HL7® FHIR®

# Questions?

Stefan Schulz:



[stefan.schulz@medunigraz.at](mailto:stefan.schulz@medunigraz.at)

<http://purl.org/steschu>

## Annotation guideline for semantic annotations of clinical narratives based on SNOMED CT and FHIR

Stefan Schulz<sup>1,2</sup>, Akhila Naz Kuppassery<sup>1</sup>, Alexander Beger<sup>1</sup>, Sareh Aghaei<sup>1</sup>, Daniel Dür<sup>1</sup>, Larissa Hammer<sup>1</sup>, Kristian Kankainen<sup>3</sup>, Markus Kreuzthaler<sup>1</sup>, NN, NN...

<sup>1</sup>Medical University of Graz, Austria  
<sup>2</sup>Averbis GmbH, Freiburg, Germany  
<sup>3</sup>Tallin University of Technology, Estonia

Version 2023-06-14

Guests are welcome! Feel free to drop comments.  
E-mail contact: stefan.schulz@medunigraz.at

In progress: consolidation of practical content for insertion into AIDAVA D4.2

<b>1. Introduction</b>	1
<b>2. Background</b>	2
2.1. Annotation strategies	2
2.2. Related work	4
<b>3. Objectives</b>	5
<b>4. Tools and resources</b>	6
4.1. Naming and graphical conventions	6
4.2. Annotation vocabularies	6
4.2.1. SNOMED CT	6
4.2.2. HL-7 FHIR	6
4.2.3. LOINC	7
4.3. INCEpTION	7
<b>5. Basic assumptions and decisions</b>	7
5.1. Conceptual model	7
5.2. Principles	7
5.3. Preferences regarding SNOMED content	10
5.3.1. Core (focus) concepts	1
5.3.2. Modifying concepts	12
5.3.3. Product concepts	12
5.3.4. Concepts from HL7 value sets	12
5.3.5. Combination of annotations	13
5.4. Concepts that require values	14
5.5. Relations, predicates, and operators	16
5.5.1. Negation, Uncertainty, clinical status and severity	16

Comment on our annotation guideline: <https://bit.ly/3X4McGC>