

Inter-Annotator-Studie basierend auf dem AIDAVA-Annotationsguide

Durchgeführt im Rahmen eines DAAD-Auslandsaufenthalts an der Med Uni Graz
Doktorandin: Andrea Riedel
Mentor: Prof. Stefan Schulz



Deutscher Akademischer Austauschdienst
German Academic Exchange Service

**Uniklinikum
Erlangen**



FAU

Friedrich-Alexander-Universität
Medizinische Fakultät



Kontext: EU-Projekt AIDAVA

- Siehe Präsentation von Stefan Schulz auf GeMTeX-Meeting Nov. 2023
- Umfangreiche Anleitungen für Annotationen klinischer Texte
 - SNOMED-Codes
 - INCEpTION
 - In zahlreichen Iterationen verfeinert
- Eigenes, benutzerfreundliches Vokabular für Relationsannotationen, abgebildet auf SNOMED und/oder FHIR-Templates
- Für Annotationen im EU-Projekt 
 - Annotation klinischer Texte zu Mamma-Ca und ischämische Herzkrankheiten
 - Ziel: „kanonische“ Annotation auf Basis der Standards SNOMED CT und FHIR
 - Annotierte Korpora zum NLP/LLM Training / Finetuning und als Benchmark
- Beispiel:

Annotation instructions	
Spans	Choose the largest sequence to which a code can be assigned to. SNOMED CT determines annotation granularity 
Overlap	Spans may be superposed and cover text that belongs to a different annotation. 
Composition	Composed meaning can be expressed By exact superposition of annotations 
Coreference	Code the literal meaning when coding anaphors (more general annotation points to a more specific annotation via anno:sameAs) 

The AIDAVA Annotation Guide (AAG) -

Instructions for semantic annotations of clinical narratives based on SNOMED CT and FHIR

Authors / Contributors:

Stefan Schulz¹, Sarah Aghavi², Akhila Naz Kappaswamy³, Alexander Bigger⁴, Daniel Dür⁵, Larissa Hemmer⁶, Kristian Karkainen⁷, Heide Albrecht⁸, Jennifer Rode⁹, Matthiasgöppel Isaradisch¹⁰, Catalina Martines-Costa¹¹, Andrea Riedel¹², Luise Modersohn¹³, Christina Lohr¹⁴, Goran Nenadic¹⁵, Warren Del-Pinto¹⁶, Ulfeng Han¹⁷, Markus Kreuzhuber¹⁸

¹Medical University of Graz, Austria, ²Universitäts GmbH, Freiburg, Germany, ³Balli University of Technology, Estonia, ⁴Chiang Mai University, Chiang Mai, Thailand, ⁵University of Murcia, Spain, ⁶University of Erlangen, Germany, ⁷Technical University of Munich, Germany, ⁸University of Leipzig, Germany, ⁹University of Manchester, UK

Version 2024 02 26



Guests are welcome! Feel free to drop comments. E-mail contact: stefan.schulz@medunigraz.at
 This site is currently subject to frequent modifications, which may affect the numbering of subsections.
 New: [Annotation Cheat Sheet \[pdf\]](#)

1. Introduction	3
2. Background	4
2.1. Annotation strategies	4
2.2. Related work	5
3. Objectives	7
4. Tools and resources	7
4.1. Annotation Tool	7
4.2. Annotation vocabularies	8
4.2.1. SNOMED CT	8
4.2.2. I9-7 FHIR	8
4.2.3. Value set for predicate annotations	8
4.2.4. Post-annotation processing and export	8
5. Basic assumptions and decisions	8
5.1. Annotation philosophy	8
5.2. General Annotation Principles	9
5.2.1. Methodology and scope of annotation	9
5.2.2. Annotation spans	10
5.2.3. Annotation vocabulary	10
5.2.4. Facts	11
5.3. Annotation symbols	12
5.3.1. Concepts	12
5.3.2. Predicates and their definitions	15

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

<https://www.aidava.eu/>

GEFÖRDERT VOM



Studienaufbau

- Zwei Annotatorinnen der Med Uni Graz hatten 6 Monate Erfahrung mit Annotationsguide, hauptsächlich an Dokumenten zu Mamma-Ca
- Für IAA-Studie Extraktion von Kurzdokumenten (Fragmente) aus dem ASSESS-CT-Korpus (sehr diverse Inhalte, teils fremdsprachige Quellen, ins Deutsche übersetzt)
- Training anhand von 10 Dokumenten, danach Besprechung der Annotationsunterschiede
- Für IAA-Experiment: 20 Dokumente
- Annotation in INCEpTION, Unterstützung durch Annotationsguide / Cheat Sheet und SNOMED-Browser
- Analyse:
 - Krippendorff's Alpha als Maß der Übereinstimmung
 - Annotationsgeschwindigkeit
 - Diskussion der Unterschiede bei der Annotation

Anamnese aus Arztbrief
Befund und Empfehlung Endokrinologieambulanz
Empfehlung aus Verlegungs-Bericht Intensiv
Medikation aus Arztbrief Innere
Untersuchungsbefund Neurologie
Verlauf aus Arztbrief Neonatologie
Konsiliarbefund Orthopädie
Gerichtsmedizinischer Autopsiebefund
Arztbrief Innere Medizin - Anamnese
Autopsiebefund Fötus
Aufnahmebefund Notfallambulanz Neurochirurgie
Konsiliarischer Befund Psychiatrie
Autopsiebefund
Aufnahmebefund Chirurgieambulanz
Medikamentenliste
Onkologische Verlaufsbeschreibung
OP-Befund der Thoraxchirurgie
Beurteilung Mamma-Chirurgie
Kontrolluntersuchung Gefäßchirurgie
Befund Tumornachsorge

Annotationsgeschwindigkeit

Konzept	Summe der Annotationszeit beider Annotatorinnen		Makroaverage der Tokenzahl	Durchsatz= (Gesamtzahl aller Tokens x 2)/ Summe der Annotationszeit beider Annotatorinnen	Makroaverage der Zeichenanzahl mit Leerzeichen	Durchsatz= (Gesamtzahl aller Zeichen x 2)/ Summe der Annotationszeit beider Annotatorinnen
	Minuten					
Training Prä-	374	Minuten	282	1,51	2047	10,95
Test Prä-	564	Minuten	594	2,11	4317	15,31

Relation	Summe der Annotationszeit beider Annotatorinnen		Makroaverage der Tokenzahl	Durchsatz= (Gesamtzahl aller Tokens x 2)/ Summe der Annotationszeit beider Annotatorinnen	Makroaverage der Zeichenanzahl mit Leerzeichen	Durchsatz= (Gesamtzahl aller Zeichen x 2)/ Summe der Annotationszeit beider Annotatorinnen
	Minuten					
Training Prä-	14	Minuten	282	40,29	2047	292,43
Test Prä-	103	Minuten	594	292,43	4317	83,83

Berechnung weiterer statistischer Kenngrößen

Qualitative Analyse der Annotationsunterschiede



GeMTeX



Qualitative Auswertung / Feedback Annotatorinnen



GeMTeX

- Nichtbeachtung bestehender Annotationsregeln
 - Codes aus exkludierten Hierarchien von SNOMED CT (z.B. Situation-Hierarchie) verwendet
 - Nichtbeachtung der Longest-Match-Regel, zu kurze Annotations-Spans, Fehlende Verankerungen in „Core“
 - Verletzung von Definitions- und Wertebereichrestriktionen bei Relationen
- Phänomene, die aus AIDAVA-Dokumenten nicht bekannt waren
 - Formulierungen von Normalbefunden
 - Umgang mit Abschnittsüberschriften und Institutionsnamen
- Probleme mit SNOMED CT
 - Unscharfe semantische Abgrenzungen ähnlicher SNOMED-Konzepte
 - Gleichwertige Varianten, denselben Inhalt unterschiedlich zu annotieren
 - „Ausgemusterte“ SNOMED-CT-Codes
- Probleme mit Annotationsguide
 - Unterschiedliche Interpretation von Annotationsregeln
 - Verlust der Übersicht durch ständige Aktualisierungen
- Eigenschaften der annotierten Textfragmente
 - Unterschiedliche Interpretationen mangels Dokumentenkontext
 - Sprachliche Ungenauigkeiten in Annotationstext, Übersetzungsfehler, komplexe semantische Strukturen
- Browsergeschwindigkeit, Browserstabilität

GEFÖRDEBT VOM



Bundesministerium
für Bildung
und Forschung

Qualitative Analyse

Terminology
Precision

SNOMED CT 

unterschiedliche
Dokumententypen und
Fragmente

Umgang mit Unbekannten

Formulare brauchen eigene
Guideline

<https://docs.google.com/document/d/1t8WZz1oTmOjTfh2fvoUauhi7KcDN70NS/edit?usp=sharing&oid=100735537636542993553&rtpof=true&sd=true>

https://docs.google.com/presentation/d/11ysZ4MQ5gvx-PfUmvsjVhFxm-8iil_2wMYa96mVuCQ/edit?usp=sharing

STEFAN

Usability für
Auswertung

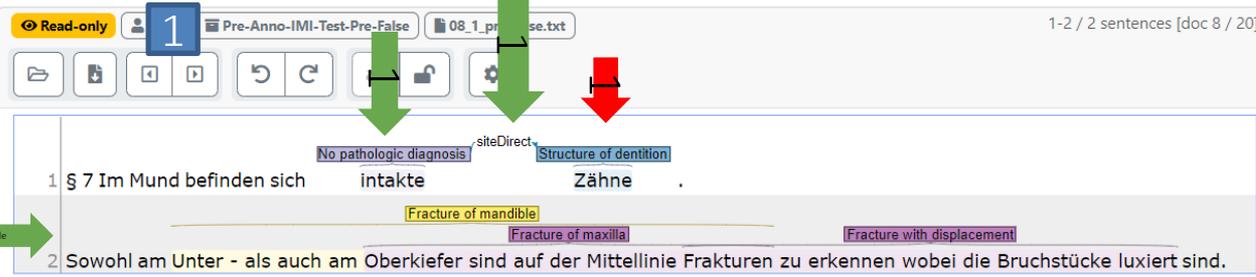
Bars lang
genug für
Lesbarkeit

STEFAN:
Umgang mit Normalbefunden
v.a. mit no pathologic
diagnosis

Read-only 1 Pre-Anno-IMI-Test-Pre-False 08_1_pre_false.txt 1-2 / 2 sentences [doc 8 / 20]

1 § 7 Im Mund befinden sich intakte Zähne .

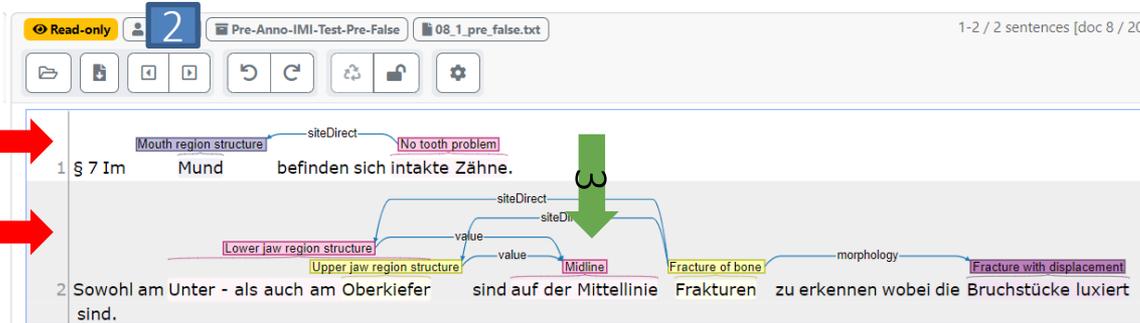
2 Sowohl am Unter - als auch am Oberkiefer sind auf der Mittellinie Frakturen zu erkennen wobei die Bruchstücke luxiert sind.



Read-only 2 Pre-Anno-IMI-Test-Pre-False 08_1_pre_false.txt 1-2 / 2 sentences [doc 8 / 20]

1 § 7 Im Mund befinden sich intakte Zähne.

2 Sowohl am Unter - als auch am Oberkiefer sind auf der Mittellinie Frakturen zu erkennen wobei die Bruchstücke luxiert sind.



1. Zeile 1: Interpretation durch Annotatorin und Guidelinekonformität:
 - a. intakt = no pathologic diagnosis → no tooth problem ist eine Situation → Guidelineverstoß
 - b. mit Relation siteDirect
 - c. Zähne = tooth structure
2. Unklarheit durch sprachliche Gestaltung des Textes: alle Zähne?
3. Zeile 2: Guidelinekonformität: Span zu detailliert und breit wie geht → wie 1, aber um Mittellinie ergänzen und beide fracture mit Relation value verknüpfen

1. **SNOMED CT Modellierungsproblem:** zeit- und aspektunabhängiges Annotieren
→ generalistisches "increase" besser, da wir Prozesszustand in zeitlichen Kontext nicht näher definieren können
2. **Guidelinekonformität:** Core concepts (wie Finding bei Beschwerden) sind zu priorisieren
3. **Neue Annotationsthemengebiete aus Patientenroutine:** Spans müsse alles relevante umfassen, Selektion von COdes nicht einfach

Read-only 1 Pre-Anno-IMI-Test-Pre-False 01_1_pre_false.txt 1-3 / 3 sentences [doc 1 / 20]

📁 📄 🔄 🔄 🗑️ 🔒 ⚙️

1 Wertigkeit ist Stadium 2 bis 3 nach NYHA .

2 Es liegt eine erhebliche Zunahme dieser Beschwerden seit Okt. '80 vor.

3 Einkaufen Haushalt und leichte körperliche Aktivitäten bereiten Schwierigkeiten.

Annotations: New York Heart Association Classification - Class III, New York Heart Association Classification - Class II, Increase, value, Multiple symptoms, beginTime, dateTime, Mild, value, Physical activity, value, Able with difficulty.

Annotations 1, 2, 3 are marked with green arrows. A red arrow 3 points to the third sentence.

Read-only 2 Pre-Anno-IMI-Test-Pre-False 01_1_pre_false.txt 1-3 / 3 sentences [doc 1 / 20]

📁 📄 🔄 🔄 🗑️ 🔒 ⚙️

1 Wertigkeit ist Stadium 2 bis 3 nach NYHA .

2 Es liegt eine erhebliche Zunahme dieser Beschwerden seit Okt. '80 vor.

3 Einkaufen Haushalt und leichte körperliche Aktivitäten bereiten Schwierigkeiten.

Annotations: New York Heart Association Classification - Class III, New York Heart Association Classification - Class II, Increasing, value, Symptom severity level, beginTime, dateTime, Personal care impairment, Impaired exercise tolerance.

Annotations 1, 2, 3 are marked with green arrows. A red arrow 3 points to the third sentence. A green arrow 3 points to the third sentence. A green arrow 3 points to the third sentence.

STEFAN:
Zeile 3

Concept Details **Expression Constraint Queries**

Concept Details

Summary | **Details** | Diagram | Expression | Refsets | Members | History | References

Stated

Parents

- Finding value (qualifier value)

Degree findings (qualifier value) ☆

SCTID: 272520006

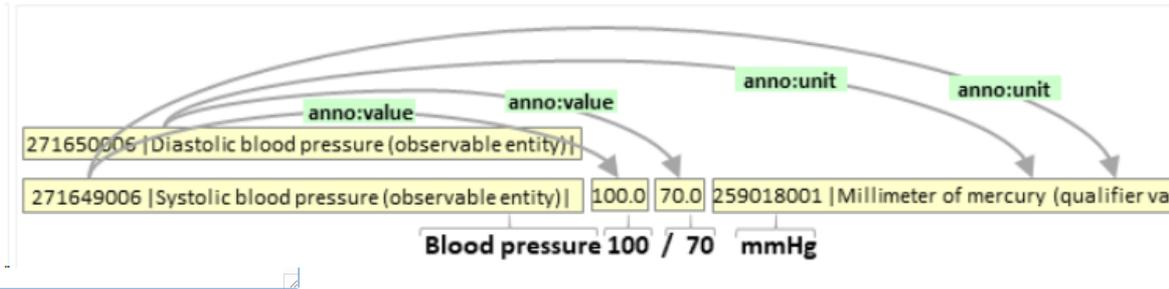
272520006 | Degree findings (qualifier value) |

en Degree findings (qualifier value)

en Degree findings

No attributes

- Highly significant (qualifier value)
- Impaired (qualifier value)
- Improved (qualifier value)
- Inaccurate (qualifier value)
- Inadequate (qualifier value)
- Inconclusive (qualifier value)
- Inconsistent (qualifier value)
- Increase (qualifier value)
- Increased (qualifier value)**
- Increasing (qualifier value)
- Indeterminate (qualifier value)
- Indifferent (qualifier value)
- Inefficient (qualifier value)
- Insufficient (qualifier value)
- Lacking (qualifier value)
- Large (qualifier value)
- Large peak (qualifier value)
- Late (qualifier value)
- Long (qualifier value)
- Long duration (qualifier value)
- Loose (qualifier value)



Schlussfolgerungen

- Es gibt nicht „die eine“ kanonische Annotation, aber bessere Approximierung möglich
- Gefahr des Ausufers des Annotationsguides
- Zu wenig klare Definitionen in SNOMED CT, zu viel ähnliche Konzepte, insbesondere im Bereich Qualifier Values
- Annotatorentaining und Annotationsunterlagen verbesserungswürdig
- Gutes Zurechtkommen mit Inception
- Gute Akzeptanz der „anno:“ Annotationsprädikate
- Spezifische Annotationsguides für spezifische Dokumententypen
 - Vergleich Vorarbeiten an Uni Murcia (Spanien):
 - Annotation von Fragebogeninhalten: erst Annotation, dann Inhalte
 - Annotation von Kliniktexten: erst Inhalte, dann Annotation
- Annotationen nur mit SNOMED CT nicht ausreichen für Kontext, Provenienz, Zeit und Zahlenwerte

BACKUP AUSWAHL?

Vorschlag: Überschriften und SNOMED Umsetzung

Kontakt Daten

Universitätsklinikum Erlangen
Medizinisches Zentrum für Informations- und Kommunikationstechnik (MIK)
Abteilung IT für Forschung & Management (IFM)
Datenintegrationszentrum (DIZ)

Universitätsstr. 22
91054 Erlangen

Tel.: +49 (9131) 85-46969

Fax: +49 (9131) 85-36799

E-Mail: andrea.riedel@uk-erlangen.de

LinkedIn: www.linkedin.com/in/andrea-riedel



Deutscher Akademischer Austauschdienst
German Academic Exchange Service

Die Arbeit wurde mit Unterstützung
eines Stipendiums im Rahmen des
IFI-Programms des DAAD ermöglicht.

Vielen Dank für Ihre
Aufmerksamkeit!