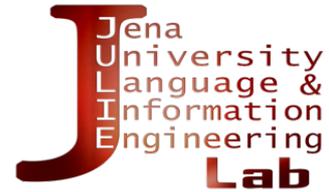




FRIEDRICH-SCHILLER-
UNIVERSITÄT
JENA



Annotationsleitlinien für deutschsprachige Medizintexte.

Teil 1: Annotation von inhaltlich kohärenten Textabschnitten (Sektionen)

Christina Lohr, Stephanie Luther, Danny Ammon,
Kutaiba Saleh, Luise Modersohn & Udo Hahn

Jena University Language & Information Engineering Lab (JULIE Lab)
Friedrich-Schiller-Universität Jena

JULIE Lab – SMITH Technischer Report 1

<https://doi.org/10.5281/zenodo.7707756>

Dezember 2022

Inhaltsverzeichnis

1	Vorwort	2
2	Einleitung	3
3	Daten	4
4	Annotationsleitlinien	4
4.1	Allgemeine Hinweise	4
4.2	Typensystem für Sektionen und Subsektionen	5
4.2.1	Präambel	5
4.2.2	Anamnese	5
4.2.3	Diagnose	5
4.2.4	Erhobene Befunde	6
4.2.5	Prozeduren und Maßnahmen	6
4.2.6	Allergien, Unverträglichkeiten, Risiken, CAVE	6
4.2.7	Medikation	6
4.2.8	Verlauf / Epikrise	7
4.2.9	Weitere empfohlene Maßnahmen	7
4.2.10	Abschließende Bemerkungen / Schlusstext	7
4.2.11	Anhang (Beilagen)	7
5	Das Annotationswerkzeug WAT-SL	8
5.1	Hinweise zur Konfiguration von WAT-SL	8
5.2	Hinweise zur Nutzung von WAT-SL	9
A	Appendix	11

1 Vorwort

Die hier dokumentierten Annotationsleitlinien für deutschsprachige Medizintexte sind im Rahmen des SMITH-Projekts¹ sowie im Projekt STAKI2B2² im Zeitraum zwischen 2017 bis 2022 entstanden.

Der vorliegende Teil 1 hat die inhaltliche Gruppierung patientenbezogener Informationen in Form von teilweise kanonisch betitelten Abschnitten (Absatz-Überschriften oder Sektionen – auch als *Section Header* bezeichnet – wie Familien- oder Patientenanamnese, Diagnose) in klinischen Dokumenten zum Gegenstand. Die verlässliche Zuordnung einzelner medizinisch relevanter Aussagen (etwa „Pat klagt über Lähmungserscheinungen der rechten Hand“) zu solchen inhaltlichen Kategorien der Abschnitte unterstützt ganz wesentlich die finale Zuordnung dieser Einzelaussagen zu grundlegenden klinischen Named-Entity-Typen (etwa Diagnosen, Befunde oder Symptome; s.a. Teil 3 der Annotationsleitlinien). Beispielsweise kann eine diagnostische Aussage im Kontext der Familienanamnese („Vater mit langjährigem DM Typ2a“) für die Patientenanamnese – und damit für den aktuellen Patienten³ – ausgeschlossen werden.

Die Zuordnung von Inhalten (etwa Diagnosen) zu den jeweiligen Kapiteln erwies sich in den uns vorgelegten klinischen Dokumenten jedoch nicht selten als nicht konsistent. So fanden sich Angaben von Medikationen im Abschnitt des Behandlungsverlaufes und nicht in der dafür vorgesehenen Sektion für Medikationsangaben.

Solche Inkonsistenzen mussten mit einer geeigneten Annotationsstrategie aufgefangen werden. Diese Beobachtung zeigt auch, dass im Allgemeinen nicht direkt von spezifischen Aussagen unter bestimmten Sektionstiteln auf gültige medizinische Aussagen zu einem bestimmten Patienten geschlossen werden kann.

Wesentlichen Einfluss auf die Entwicklung des hier beschriebenen Annotationsschemas hatten die Vorgaben, die im Kontext der HL7-Spezifikationen in der *Clinical Document Architecture*⁴ (CDA) formuliert sind. Diese mussten aufgrund der Abweichungen realer klinischer Dokumente von diesen Standards jedoch adaptiert werden.

¹Das SMITH-Projekt (<https://www.smith.care>) wurde im Rahmen der Medizininformatik-Initiative (MII – <https://www.medizininformatik-initiative.de>) vom Bundesministerium für Bildung und Forschung (BMBF) gefördert (Fördernummer: 01ZZ1803G).

²Das Projekt STAKI2B2 (<https://gepris.dfg.de/gepris/projekt/315098900>) wurde von der Deutschen Forschungsgemeinschaft unter dem Förderkennzeichen HA 2097/8-1 gefördert.

³Wir verwenden dem sprachlichen Usus der Mehrheit der Sprechergemeinschaft des Deutschen und linguistischen Argumenten folgend das generische Maskulinum zur Referenzierung aller Geschlechter.

⁴<https://hl7.de/themen/hl7-cda-clinical-document-architecture/>

2 Einleitung

Die Einteilung von Lang- oder Volltexten in makrostrukturelle, inhaltlich kohärente Einheiten (sog. „Sektionen“) ist eine elementare Voraussetzung für die automatisierte Verarbeitung von medizinischen Texten. Durch diese Einteilung wird der Kontext von Aussagen, die mit Verfahren der Informationsextraktion automatisch identifiziert werden sollen, näher charakterisiert. Dieser Kontext wiederum ist häufig nötig, um die Bedeutung von Aussagen zu disambiguieren – sowohl für menschliche als auch für maschinelle Leser. Wissen über diesen Kontext kann durchaus explizit im Text stehen, aber durch Verkürzungen oder Verletzung sprachlicher Regeln so schwer zu erkennen sein, sodass alternative Interpretationshilfen nötig werden – hier also die Zugehörigkeit zu einer Sektion.

Im Zuge von Standardisierungsbestrebungen im Gesundheitswesen wurden Regelungen entwickelt, um die inhaltliche Gliederung von Arztbriefen zu vereinheitlichen. Der HL7-Standard „Clinical Document Architecture“ (CDA) ist ein XML-basierter Markup-Standard für den Aufbau klinischer Dokumente sowie deren Austausch (1; 2). In der national für Deutschland adaptierten CDA-Spezifikation „Arztbrief Plus“ sowie deren Definition „CDA Section Level Templates“⁵ sind Regelungen zum strukturierten Verfassen von Arztbriefen, einschließlich Überschriften von Unterabschnitten und dem Inhalt bestimmter Abschnitte enthalten. Zwar waren die CDA-Standards zum Zeitpunkt der Annotationsarbeiten (2017-2018) rechtlich noch nicht verbindlich, setzen sich aber im klinischen Alltag immer stärker durch. Vor der Entwicklung der CDA-Richtlinien war die Wahl von Abschnitten sowie deren Überschriften mehr oder weniger dem verfassenden Arzt überlassen oder durch einrichtungsspezifische Vorschriften bzw. Empfehlungen vorgegeben.⁶

Die im Folgenden dokumentierten Annotationsleitlinien dienten unserem Annotatorenteam als Grundlage für die Kennzeichnung sektionaler Metadaten in den klinischen Dokumenten. Sie wurden im Verlauf von vier Vorannotationsrunden erarbeitet, in denen unter anderem ermittelt wurde, mit welchen Kategorien das Gesamtkorpus zu annotieren sei. Ein **Annotat** bezeichnet hier ein inhaltlich markiertes Textstück (ein Einzelwort oder eine Spanne von Wörtern) in einem Arztbrief.

Durch die Bereitstellung dieser Leitlinien soll es der Forschungsgemeinschaft ermöglicht werden, einerseits unsere Daten zu reproduzieren, andererseits unsere Vorschläge in zukünftigen Projekten zu nutzen bzw. weiterzuentwickeln.

Der Ergebnisse, die während der Entwicklung der hier vorliegenden Annotationsleitlinie entstanden sind, wurden wie folgt veröffentlicht:

(3) Lohr et al., *CDA-Compliant Section Annotation of German-Language Discharge Summaries: Guideline Development, Annotation Campaign, Section Classification*; AMIA 2018 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6371337/>).

Während der Annotationsarbeiten wurde das Annotationswerkzeug WAT-SL weiterentwickelt und wie folgt veröffentlicht:

(4) Lohr et al., *Continuous Quality Control and Advanced Text Segment Annotation with WAT-SL 2.0*; LAW@ACL 2019 (<https://aclanthology.org/W19-4025/>)

⁵https://wiki.hl7.de/index.php?title=IG:Arztbrief_Plus#CDA_Section_Level_Templates

⁶Abgesehen von einigen teils kodifizierten, meist aber informellen und nicht schriftlich festgelegten Regeln bezüglich der Abfolge und des Inhalts von Sektionen.

3 Daten

Die Annotationen, für die die hier dokumentierten Leitlinien entwickelt wurden, wurden an realen klinischen Arztbriefen durchgeführt. Diese Briefe sind Teil des 3000PA-Textkorpus (5). Das Korpus wurde im Rahmen der Pilotphase des SMITH-Projekts an den Standorten Jena, Leipzig und Aachen nach folgenden Kriterien zusammengestellt:

- Alle Briefe entstanden zwischen 2010 bis 2015.
- Alle in den Briefen beschriebenen Patienten waren mindestens fünf Tage in einem der Universitätsklinikum Jena, Leipzig oder Aachen stationär in Behandlung.
- Die Behandlung fand auf einer intensivmedizinischen oder internistischen Station statt.
- Alle Patienten waren zum Zeitpunkt der Datenerhebung im September 2016 bereits verstorben.

Das Jenaer Segment von 3000PA besteht insgesamt aus 1106 deutschsprachigen Arztbriefen, die sich aufteilen in:

- 383 Kurzberichte,
- 103 Verlegungsbriefe und
- 620 Entlassbriefe.

4 Annotationsleitlinien

4.1 Allgemeine Hinweise

Durch die CDA-Konventionen sind verschiedene Kategorien vorgegeben, in die Arztbriefe unterteilt werden sollen. Aus diesen wurden die für die Sektionsannotation relevanten 10 Kategorien ausgewählt: *Präambel, Erhobene Befunde, Anamnese, Diagnose, Prozeduren und Maßnahmen, CAVE, Medikation, Verlauf, Weitere empfohlene Maßnahmen, Schlusstext.*

Während der Probeannotationen experimentierten wir auch mit der Kategorie „Mix“ als Bestandteil des Typen-Systems. Diese wurde aber aus der finalen Annotationsrunde (und damit aus dieser Dokumentation) wegen ihrer semantischen Unschärfe wieder ausgeschlossen. Da die hier definierten Typen von Sektionen nicht zwingend den Überschriften und Einteilungen der vorliegenden Arztbriefe entsprachen, entschlossen wir uns, auf Satzebene zu annotieren. Für die hier vorliegende Aufgabe gilt demzufolge, dass ein Annotationsobjekt aus einem Satz besteht.

Für die finale Annotation bestand die Aufgabe darin, die unterschiedlichen Dokumenttypen satzweise jeweils einer der o.g. 10 Kategorien zuzuordnen. Hierfür erhielt jeder Annotator 182 bzw. 183 Dokumente, welche sowohl Kurzberichte, Verlegungsbriefe als auch Entlassbriefe sein konnten.

Das grundsätzliche Vorgehen besteht darin, dass ein Annotator das Symbol am Ende eines Satzes anklickt und auswählt, welcher der Kategorien bzw. Sektionstypen der vorangehende Satz zuzuordnen ist.

4.2 Typensystem für Sektionen und Subsektionen

In den folgenden Abschnitten sind die 10 grundlegenden Sektionstypen sowie ggf. ihnen zugeordnete spezifischere Subtypen beschrieben. Der Gesamtaufbau des Typensystems wurde in Rücksprache mit Kollegen des Universitätsklinikums Jena aus dem CDA-Standard „Arztbrief Plus“ mit der Definition der „CDA Section Level Templates“⁷ abgeleitet.

Es gilt bei der Annotationsentscheidung der Grundsatz, dass die *am besten passende* Kategorie auszuwählen ist. Dies gilt sowohl dann, wenn ein zu annotierendes Element (Satz) mehreren Kategorien zugeordnet werden könnte (aber eine Kategorie *passender* erscheint – Typenambiguität), als auch dann, wenn ein Satz passender einer Subsektion als der ihr übergeordneten Sektion (z.B. Anamnese – Familienanamnese) zugeordnet werden könnte (Typenspezifität). Eine Zusammenfassung des Typensystems für Sektionen ist unter Appendix A aufgelistet.

4.2.1 Präambel

Dieser Sektionstyp basiert auf der Kategorie **Anrede**, wie sie im CDA-Standard aufgeführt ist, und enthält die „allgemeinen einleitenden Sätze eines Dokuments, z.B. eines Arztbriefes oder eines Befund-Dokuments“. Hierzu gehören neben der Anschrift und Anrede des Empfängers (z.B. Anschrift des Hausarztes) die ersten einleitenden Sätze (z.B. „*Sehr geehrte Frau Kollegin, ...*“ / „*Sehr geehrter Herr Kollege, ...*“) sowie die erste Nennung des Patienten mit den jeweiligen Stammdaten (i.d.R. Name und Geburtsdatum).

4.2.2 Anamnese

- **Patientenanamnese**
- **Familienanamnese**

In der Sektion **Anamnese** sollen die konkrete (medizinische) Fragestellung bzw. der Grund für eine Überweisung aufgeführt werden. Des Weiteren soll der bei Einweisung aktuelle medizinische Zustand des Patienten ebenso wie vorherige Erkrankungen (Subsektion **Patientenanamnese**), Gesundheitsprobleme sowie ggf. eine Zusammenfassung medizinischer Gegebenheiten im familiären Umfeld desselben (Subsektion **Familienanamnese**) erläutert werden.

4.2.3 Diagnose

- **Frühere Diagnosen**
- **Aktuelle Diagnose**

Unter dem Sektionstyp **Diagnose** werden alle Diagnosen aufgeführt, die vor dem beschriebenen Krankenhausaufenthalt gestellt wurden (Subsektion **Frühere Diagnosen**) oder im Verlauf der Entlassuntersuchung (Subsektion **Aktuelle Diagnose**) gestellt werden. Können diese Unterscheidungen nicht getroffen werden, wird der allgemeinere Sektionstyp **Diagnose** gewählt. Zu dieser Kategorie gehören auch ICD-Diagnosen und TNM-Klassifikationen.

⁷http://wiki.hl7.de/index.php?title=IG:Arztbrief_Plus

4.2.4 Erhobene Befunde

Dieser Sektionstyp umfasst während des Krankenhausaufenthaltes erhobene Befunde aus der körperlichen Untersuchung, Bildgebung und auch Labordaten — jedoch nur solche, die im Fließtext vorkommen (keine Labordaten, die im Anhang des Arztbriefes tabellarisch aufgeführt werden). Beispiele:

- „*Pulmo: Basal diskrete Rgs*“,
- „*Abdomen: weich, Peristaltik: +++*“ oder
- „*Haut blass, seborrhoisches Ekzem, Schleimhäute blass, Hautturgor herabgesetzt*“.

4.2.5 Prozeduren und Maßnahmen

Dieser Sektionstyp deckt insgesamt eine „Kurzbeschreibung sämtlicher, während des Aufenthalts durchgeführter Maßnahmen, darunter OPs, Eingriffe und sonstige Maßnahmen“ ab und enthält u.a. Informationen über:

- „Fachspezifische Eingriffe
- Operationen
- Strahlentherapie
- Lichttherapie
- Psychiatrische Therapie“

Erwähnungen, die in den CDA-Standard „Heil- und Hilfsmittel“ eingeordnet werden können, fallen ebenfalls in diese Kategorie.

4.2.6 Allergien, Unverträglichkeiten, Risiken, CAVE

Dieser Sektionstyp bezieht sich auf allfällige Hinweise zu Risikofaktoren und Allergien eines Patienten. Darin sind Beschreibungen der Allergien, Unverträglichkeiten und Risiken sowie deren beobachtete Nebenwirkungen als auch sonstige Risiken enthalten, z.B.: „*Penicillinallergie*“.

4.2.7 Medikation

- **Medikation bei Einweisung**
- **Medikation während des Aufenthaltes**
- **Medikation bei Entlassung**

Dieser Sektionstyp zielt auf alle Medikationsangaben, die innerhalb des Arztbriefes erwähnt werden (inklusive Angaben zur Dosierung). Auch verabreichte Impfungen sind dieser Kategorie zuzuordnen. Für Angaben zu Medikationen, bei denen eindeutig geklärt ist, dass diese *vor* dem Aufenthalt genommen wurde, ist **Medikation bei Einweisung** zu wählen. Des Weiteren kann erwähnt sein, dass die Medikation *bei* Entlassung verabreicht wurde; hier ist der Subsektionstyp **Medikation bei Entlassung** auszuwählen. Wird gesondert erwähnt, dass die Medikation *während* des Krankenhausaufenthaltes des Patienten gegeben wurde, ist der Subsektionstyp **Medikation während des Aufenthaltes** zu kodieren. Können diese Unterscheidungen nicht getroffen werden, wird der allgemeinere Sektionstyp **Medikation** gewählt.

4.2.8 Verlauf / Epikrise

Der Sektionstyp **Verlauf / Epikrise** ist an die CDA-Sektion *Epikrise / Zusammenfassung des Aufenthalts* angelehnt. Sätze, die einen speziellen zusammenfassenden Rückblick, eine Interpretation des Krankenhausgeschehens sowie der veranlassten Therapie beschreiben, sollen damit annotiert werden. Beinhalten Sätze Beschreibungen zu Befund und Therapie, so ist dies ebenfalls mit **Verlauf / Epikrise** zu kennzeichnen.

4.2.9 Weitere empfohlene Maßnahmen

Dieser Sektionstyp beinhaltet Empfehlungen für weitere, noch durchzuführende Maßnahmen, z.B. durch den weiterbehandelnden Arzt.

4.2.10 Abschließende Bemerkungen / Schlusstext

Der Sektionstyp **Schlusstext** ist den am Ende des Arztbriefes formulierten Freitextpassagen zuzuordnen, die oft mit Verabschiedungsformeln und Namensnennungen des Briefautors verknüpft sind (z.B. „*Mit kollegialen Grüßen ...*“) oder mit für die berichtende Einrichtung besonders typischen Floskeln.

4.2.11 Anhang (Beilagen)

Dieser Sektionstyp umfasst alle sonstigen Beilagen und Anhänge, außer denjenigen Dokumenten, die in „Patientenverfügungen und andere juristische Dokumente“ angegeben sind. Der Anhang ist entweder direkt im Dokument enthalten oder wird durch einen Hinweis auf ein externes Dokument referenziert.

5 Das Annotationswerkzeug WAT-SL

Die folgenden Abschnitte beschreiben die Installation bzw. Konfiguration sowie die Verwendung des WAT-SL-Annotationswerkzeugs für die manuelle Sektionsannotation in deutschsprachigen Arztbriefen.

5.1 Hinweise zur Konfiguration von WAT-SL

Für die manuelle Annotation von Textabschnitten wird das WEB ANNOTATION TOOL FOR SEGMENT LABELING – WAT-SL 2.0^{8 9} verwendet.(4; 6) WAT-SL ähnelt sowohl im Aufbau als auch in seiner Funktionalität dem weitverbreiteten Werkzeug BRAT¹⁰.(7) Jedoch ist BRAT für die Annotation längerer Textpassagen, die über mehrere Zeilen hinweglaufen, instabil, sodass wir alternativ das dafür geeignetere Werkzeug WAT-SL an die Annotation von längeren Textabschnitten für klinische Berichte konfiguriert haben. WAT-SL wird wie BRAT ebenfalls ausschließlich über einen Internet-Browser verwendet und kann entweder lokal oder auf einem Server installiert werden.

Vorab muss WAT-SL konfiguriert werden und für jeden Benutzer jeweils ein Profil eingerichtet werden. Die notwendigen Konfigurations-Dateien werden unter <https://doi.org/10.5281/zenodo.7707756> bereitgestellt.

Textdateien müssen vorab aufbereitet sein. Jedes Element, das eine zu annotierende Sektion darstellt, muss durch ein Newline-Zeichen ("\n") voneinander abgetrennt sein. Werden einzelne Sätze annotiert, so besteht eine einzelne Zeile aus einem einzelnen Satz.

Die Annotationsdateien werden in einem csv-Format ausgegeben und ebenso in einem BRAT-kompatiblen Format.

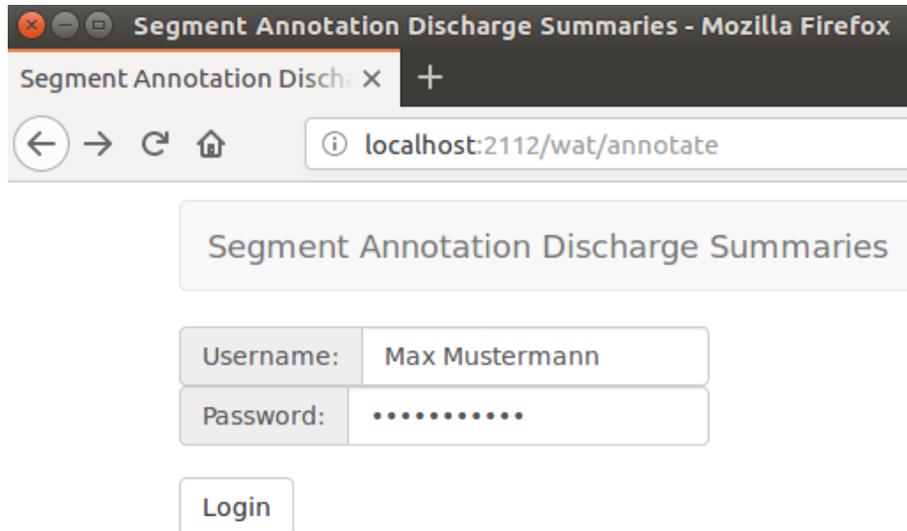
⁸<https://github.com/webis-de/wat>

⁹<https://doi.org/10.5281/zenodo.3354030>

¹⁰<http://brat.nlplab.org/index.html>

5.2 Hinweise zur Nutzung von WAT-SL

Vor der Benutzung muss sich jeder Annotator mit einem Benutzernamen und Passwort am Annotationssystem anmelden (s. Abb. 1).



Segment Annotation Discharge Summaries

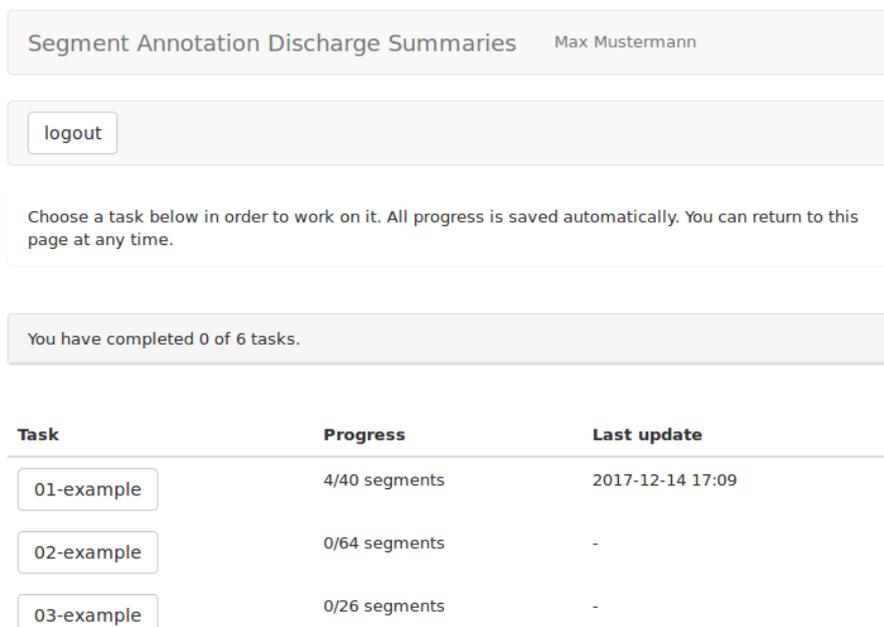
Username: Max Mustermann

Password:

Login

Abbildung 1: Anmeldung am System WAT-SL zur Annotation

Anschließend ist die zu bearbeitende Aufgabe auszuwählen (s. Abb. 2).



Segment Annotation Discharge Summaries Max Mustermann

logout

Choose a task below in order to work on it. All progress is saved automatically. You can return to this page at any time.

You have completed 0 of 6 tasks.

Task	Progress	Last update
01-example	4/40 segments	2017-12-14 17:09
02-example	0/64 segments	-
03-example	0/26 segments	-

Abbildung 2: Auswahl der Annotationsaufgaben im System WAT-SL

Durch einen Klick auf das Symbol (roter Kreis mit Kreuz), das einem Satz folgt, wird ein Menü zur Auswahl von Annotationen aufgerufen (s. Abb. 3).

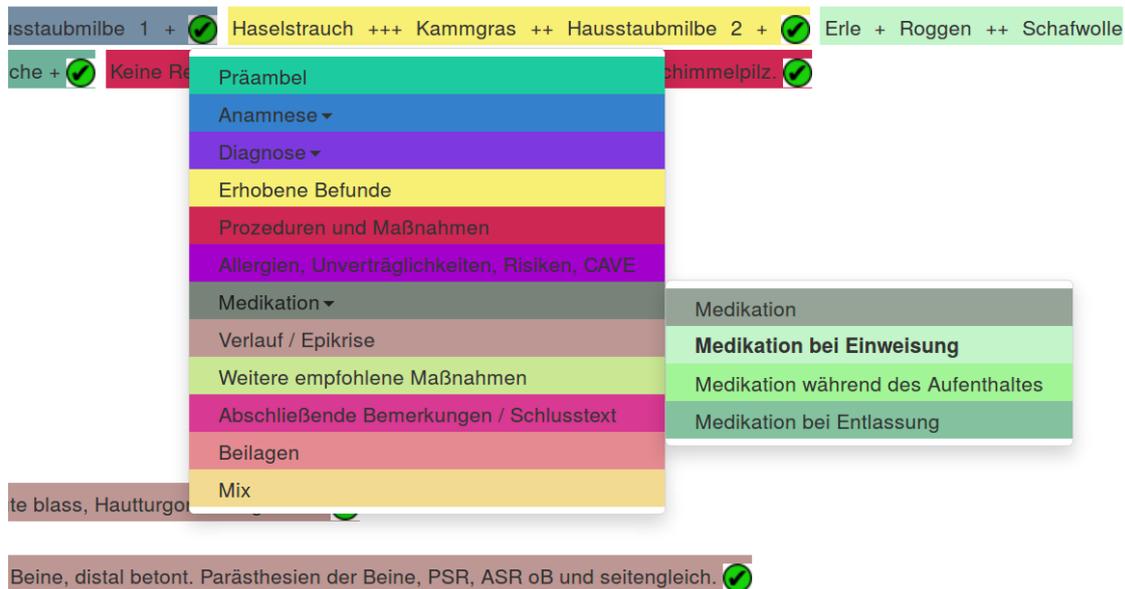


Abbildung 3: Auswahl einer Kategorie

Für die bessere Unterscheidbarkeit und ein komfortableres Arbeiten wurden die einzelnen Kategorien mit Farben gekennzeichnet. Weiterhin wird diejenige Kategorie, welche zuletzt ausgewählt wurde, im Auswahlmenü fett markiert.

Für jede Kategorie wurde ein Short-Cut angelegt, sodass die Auswahl auch durch Tastaturbelegungen angesteuert werden kann, siehe dazu Appendix A.

A Appendix

Typ	Beschreibung	Short-Cut
Präambel	Anrede, allgemeine einleitende Sätze, Stammdaten des Patienten	Ä
Anamnese	Angaben zur Anamnese	–
Patientenanamnese	Angaben zur Anamnese des Patienten; Fragestellung bzw. Grund der Ein- oder Überweisung, aktueller medizinischer Zustand, Vorerkrankungen	A
Familienanamnese	Diagnosen, die von Verwandten des Patienten bekannt sind	F
Diagnose	Angaben zu Diagnosen; ICD-Diagnosen und TNM-Klassifikationen	D
Frühere Diagnose	speziell gekennzeichnete Diagnose, im Verlauf der Aufnahmeuntersuchung gestellt	Ü
Entlassdiagnose / aktuelle Diagnose	Diagnose, mit der der Patient entlassen wurde	K
Erhobene Befunde	während des Krankenhausaufenthaltes erhobene Befunde	B
Prozeduren und Maßnahmen	während des Aufenthalts durchgeführte Maßnahmen, darunter OPs, Eingriffe und sonstige Maßnahmen: fachspezifische Eingriffe, Operationen, Strahlentherapie, Lichttherapie, Psychiatrische Therapie	P
Allergien, Unverträglichkeiten, Risiken, CAVE	Risikofaktoren, Allergien, Unverträglichkeiten, Nebenwirkungen	C
Medikation	Angaben zu eingenommenen und verabreichten Medikamenten	M
Medikation bei Einweisung	bei Aufnahme des Patienten erhobene (bestehende) Medikation	1
Medikation während des Aufenthaltes	erhaltene Medikamente des im Brief beschriebenen Krankenhausaufenthaltes	2
Medikation bei Entlassung	Medikation bei Entlassung des Patienten	3
Verlauf / Epikrise	zusammenfassender Rückblick aus Interpretation des Krankengeschehens und veranlasste Therapien, Beschreibungen zu Verlauf des Aufenthaltes	V

Typ	Beschreibung	Short-Cut
Weitere empfohlene Maßnahmen	Empfehlungen für weitere, noch durchzuführende Maßnahmen	W
Abschließende Bemerkungen / Schlusstext	am Ende des Briefes formulierter Freitext	S
Beilagen	alle sonstigen Beilagen/Anhänge, außer Patientenverfügung und andere juristische Dokumente	L

Referenzen

- [1] Keith W Boone. *The CDA™ Book*. Springer Science & Business Media, 2011.
- [2] Tim Benson and Grahame Grieve. CDA — Clinical Document Architecture. In Tim Benson and Grahame Grieve, editors, *Principles of Health Interoperability: SNOMED CT, HL7 and FHIR*, Health Information Technology Standards (HITS), pages 283–301. Springer, Cham, Switzerland, 3rd edition, 2016.
- [3] Christina Lohr, Stephanie Luther, Franz Matthies, Luise Modersohn, Danny Ammon, Kuntaiba Saleh, Andreas Henkel, Michael Kiehntopf, and Udo Hahn. CDA-compliant section annotation of German-language discharge summaries: guideline development, annotation campaign, section classification. In *AMIA 2018 — Proceedings of the 2018 Annual Symposium of the American Medical Informatics Association. Data, Technology, and Innovation for Better Health. San Francisco, California, USA, November 3-7, 2018*, pages 770–779, 2018.
- [4] Christina Lohr, Johannes Kiesel, Stephanie Luther, Johannes Hellrich, Benno Stein, and Udo Hahn. Continuous annotation quality control, support for hierarchically structured label sets and long-segment annotation with WAT-SL 2.0. In Annemarie Friedrich and Deniz Zeyrek, editors, *LAW XIII 2019 — Proceedings of the 13th Linguistic Annotation Workshop @ ACL 2019. Florence, Italy, August 1, 2019*, pages 215–219, Stroudsburg/PA, 2019. Association for Computational Linguistics (ACL).
- [5] Udo Hahn, Franz Matthies, Christina Lohr, and Markus Löffler. 3000PA: towards a national reference corpus of German clinical language. In Adrien Ugon, Daniel Karlsson, Gunnar O. Klein, and Anne Moen, editors, *Building Continents of Knowledge in Oceans of Data: The Future of Co-Created eHealth - Proceedings of MIE 2018, Medical Informatics Europe, Gothenburg, Sweden, April 24-26, 2018*, volume 247 of *Studies in Health Technology and Informatics*, pages 26–30. IOS Press, 2018.
- [6] Johannes Kiesel, Henning Wachsmuth, Khalid Al-Khatib, and Benno Stein. WAT-SL: a customizable Web annotation tool for segment labeling. In Anselmo Peñas and André F. T. Martins, editors, *EACL 2017 — Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Software Demonstrations. Valencia, Spain, April 5-6, 2017*, pages 13–16, Stroudsburg/PA, 2017. Association for Computational Linguistics (ACL).
- [7] Pontus Stenetorp, Sampo Pyysalo, Goran Topić, Tomoko Ohta, Sophia Ananiadou, and Jun’ichi Tsujii. BRAT: A Web-based tool for NLP-assisted text annotation. In Frédérique Segond, editor, *EACL 2012 — Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics: Demonstrations. Avignon, France, April 25-26, 2012*, pages 102–107, Stroudsburg/PA, 2012. Association for Computational Linguistics (ACL).