

JULIELAB@SMITH

Annotation von Medikationsinformationen in deutschsprachigen Entlassbriefen

Franz Matthies

Lehrstuhl für Computerlinguistik
Institut für Germanistische Sprachwissenschaft
Friedrich-Schiller-Universität Jena
www.julielab.de

29. Juni 2017

Überblick

- 1 Aufgabe(n)
- 2 Annotation
 - Begriffe
 - Korpus
 - Annotationsvorgang
- 3 Automatische Extraktion
- 4 Auswertung
 - Centroid
 - IAA & Systemperformanz

Überblick

- 1 Aufgabe(n)
- 2 Annotation
 - Begriffe
 - Korpus
 - Annotationsvorgang
- 3 Automatische Extraktion
- 4 Auswertung
 - Centroid
 - IAA & Systemperformanz

Aufgabe(n)

- Manuelle Annotation von Medikationsinformationen
 - dafür Erstellung von Guidelines
(Orientierung an I2B2 Guidelines)
- Erstellung/Anpassung eines automatischen Medikationsextraktionssystems
(aufbauend auf MEDXN System)
- Auswertung der Annotationsgüte und des Extraktionssystems

Überblick

1 Aufgabe(n)

2 Annotation

- Begriffe
- Korpus
- Annotationsvorgang

3 Automatische Extraktion

4 Auswertung

- Centroid
- IAA & Systemperformanz

Begriffe

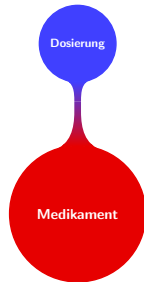
Medikation



Substanzen, die einge-
nommen oder zugeführt
wurden bzw. werden
(rezeptpflichtig, frei-
verkäuflich, biologisch;
auch Kontrastmittel u.ä.)

Begriffe

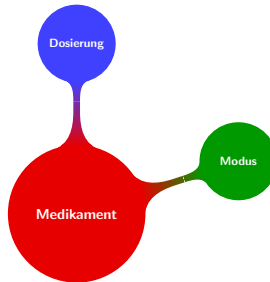
Medikation



Angaben, wieviel eines Medikaments per Verabreichung genommen wird (100ml, 1 Tablette, 4 IE, etc.)

Begriffe

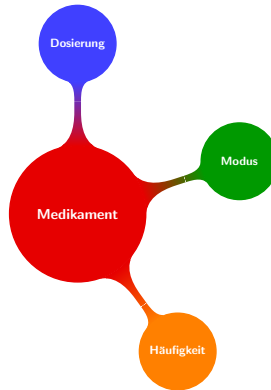
Medikation



Auf welche Weise wurde
das Medikament verab-
reicht
(oral, intravenös, etc.)

Begriffe

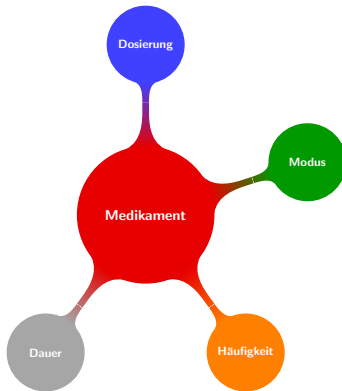
Medikation



Numerische & alpha-
betische Angaben der
Häufigkeit einer Einnah-
me/ Vergabe
(1x, einmal im Monat,
nach Bedarf, 1-0-0, nach
dem Essen, etc.)

Begriffe

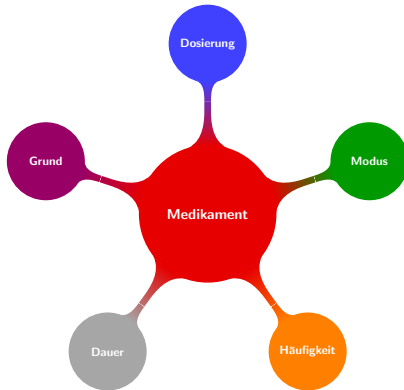
Medikation



Konkrete und vage Angaben für den Zeitraum bzw. wie lange das Medikament eingenommen/ verabreicht wurde
(für 10 Tage, seit 11/2010, 10 Packungen, bis die Symptome verschwinden)

Begriffe

Medikation



Medizinischer Grund bzw. der Anlass für die Verschreibung/ Einnahme des Medikaments
(Fieber, Kopfschmerzen, Diabetes)

Korpus: Angaben

- 3 (Standorte) * 1000 Patientenakten (Entlass- & Verlegungsbriefe)
- Nicht anonymisiert oder pseudonymisiert; Patienten alle schon verstorben
- *txt*- oder wenigstens *doc(x)*- Format

Korpus: Fließtext

- 23 Anamnese und Verlauf:
- 24 Herr XXX wurde am XXX mit Allgemeinzustandsverschlechterung bei persistierenden massiven Diarrhoen (bis zu mehreren Litern Wasser pro Stuhlgang), Übelkeit und Erbrechen mit allgemeinem Schwächegefühl und Kopfschmerzen von der hiesigen zentralen Notaufnahme übernommen.
- 25 Er wurde im Vorfeld XXX mehrfach wegen der gleichen Symptomatik vorstellig.
- 26 In den Stuhlproben konnten jeweils kein Erreger nachgewiesen werden.
- 27 Schließlich wurde im Rahmen eines stationären Aufenthaltes XXX in unserer Klinik für Gastroenterologie die Diagnose eines neuroendokrinen Tumors im Pankreasschwanz mit histologisch gesicherten Lebermetastasen gestellt werden.
- 28 Es wurde die Therapie mit Octreotid begonnen.
- 29 Der Verdacht auf MEN1 bei Hyperkalziämie und hochnormalem Parathormon mit Hyperparathyreoidismus konnte nicht sicher ausgeschlossen werden.
- 30 Ein Hypophysentumor kam im Rahmen der cerebralen Bildgebung nicht zur Darstellung.
- 31 Der Patient berichtet, dass er XXX einmalig Octreotid i.m. in der hiesigen gastroenterologischen Ambulanz erhalten habe und seither kein Octreotid verabreicht bekommen hätte (auch nicht s.c.), da bei folgender geplanter monatlicher Vorstellung zur i.m.-Gabe die Gabe pausiert worden sei bei Infektion im Perianalbereich, welche im Krankenhaus Wurzen operiert worden sei und wohl noch werden soll.
- 32 Darunter habe sich die Symptomatik mit massiven Diarrhoe, Bauchschmerzen, Übelkeit und Erbrechen verschlechtert.
- 33 Er habe kaum Flüssigkeit oder Nahrung aufnehmen können.
- 34 Laborchemisch imponierte schließlich eine metabolische Azidose bei akutem Nierenversagen und eine nur leichte Hyperkalziämie.
- 35 Herr XXX wurde zur weiteren Therapie und Überwachung übernommen.
- 36 Unter Volumengabe, symptomatischer Therapie und s.c.
- 37 Octreotidgabe konnte eine schnelle Besserung der Metabolik erreicht werden mit stabiler Diurese, die Retentionswerte fielen in den Normbereich.
- 38 Ein Dialyseverfahren erwies sich als nicht notwendig.
- 39 Die Symptomatik der Diarrhoe besserte sich deutlich.
- 40 Wir verlegen daher den Pateinten in deutlich gebessertem Zustand auf Ihre Station und danken für die prompte Übernahme.

Korpus: Fließtext – annotiert

- 23 Anamnese und Verlauf:
- 24 Herr XXX wurde am XXX mit Allgemeinzustandsverschlechterung bei persistierenden massiven Diarrhoen (bis zu mehreren Litern Wasser pro Stuhlgang), Übelkeit und Erbrechen mit allgemeinem Schwächegefühl und Kopfschmerzen von der hiesigen zentralen Notaufnahme übernommen.
- 25 Er wurde im Vorfeld XXX mehrfach wegen der gleichen Symptomatik vorstellig.
- 26 In den Stuhlproben konnten jeweils kein Erreger nachgewiesen werden.
- 27 Schließlich wurde im Rahmen eines stationären Aufenthaltes XXX in unserer Klinik für Gastroenterologie die Diagnose eines neuroendokrinen Tumors im Pankreasschwanz mit histologisch gesicherten Lebermetastasen gestellt werden.
- 28 Es wurde die Therapie mit **Medikament** Octreotid begonnen.
- 29 Der Verdacht auf MEN1 bei Hyperkalziämie und hochnormalem Parathormon mit Hyperparathyreoidismus konnte nicht sicher ausgeschlossen werden.
- 30 Ein Hypophysentumor kam im Rahmen der cerebralen Bildgebung nicht zur Darstellung.
- 31 Der Patient berichtet, dass er XXX einmalig **Medikament** Octreotid **Modus** i.m. in der hiesigen gastroenterologischen Ambulanz erhalten habe und seither kein Octreotid verabreicht bekommen hätte (auch nicht s.c.), da bei folgender geplanter monatlicher Vorstellung zur i.m.-Gabe die Gabe pausiert worden sei bei Infektion im Perianalbereich, welche im Krankenhaus Wurzen operiert worden sei und wohl noch werden soll.
- 32 Darunter habe sich die Symptomatik mit massiven Diarrhoe, Bauchschmerzen, Übelkeit und Erbrechen verschlechtert.
- 33 Er habe kaum Flüssigkeit oder Nahrung aufnehmen können.
- 34 Laborchemisch imponierte schließlich eine metabolische Azidose bei akutem Nierenversagen und eine nur leichte Hyperkalziämie.
- 35 Herr XXX wurde zur weiteren Therapie und Überwachung übernommen.
- 36 Unter Volumengabe, symptomatischer Therapie und **Modus** s.c.
- 37 **Medikament** Octreotidgabe konnte eine schnelle Besserung der Metabolik erreicht werden mit stabiler Diurese, die Retentionswerte fielen in den Normbereich.
- 38 Ein Dialyseverfahren erwies sich als nicht notwendig.
- 39 Die Symptomatik der Diarrhoe besserte sich deutlich.
- 40 Wir verlegen daher den Patienten in deutlich gebessertem Zustand auf Ihre Station und danken für die prompte Übernahme.

Korpus: Medikationsliste

43	Medikamentöse Therapie bei Verlegung:
44	i.v.
45	Pantozol 40 mg 1-0-1
46	Granisetron 1 mg 1-0-0
47	Novamin 1000 mg 1-1-1
48	s.c.:
49	Octreotid 100 µg 1-1-1
50	Innohep 3500 IE 0-0-1

Korpus: Medikationsliste – annotiert



Annotationsvorgang (Jena)

- Semi-zufällige Auswahl von fünfzehn Dokumenten (Sicherstellung dass überhaupt “Medikamente” enthalten sind)
- Drei Iterationen über je fünf Dokumente durch die Jenaer Annotatorinnen mit Feedbackdiskussionsrunden nach jedem Block
→ Schulung + Guidelines verbessern
- Im Anschluss zufällige Aufteilung der Gesamtmenge an Dokumenten; inklusive 52 gleicher Dokumente für alle (zur Gütemessung)

Annotationsvorgang: Daten

	UKJ	UKA	UKL	Bemerkung
Rechtliche Rahmenbedingungen	+	+	+	
Zugang zu Räumlichkeiten (UKJ)	+	o	o	12.12.2016
Technischer Zugang zu Dokumenten vor Ort (Infrastruktur verfügbar)	+ (13.12.16)	+ (im Dez 16)	+ (30.01.17)	
Technischer Zugang zu Dokumenten aus Jena (Infrastruktur verfügbar)	+ (13.12.16)	o (19.01.17)*	o (09.03.17)	*Zugang UKA sehr fragil
Annotations-Guidelines (V 0.9) verfügbar	+ (06.01.17)	+ (23.01.17)*	+ (23.01.17)	**UKA: Guideline für Annotatorin bis 31.1. nicht verfügbar
Start Annotation	+ (11.01.17)	o (08.03.17)	+ (30.01.17)**	**UKL: Annotatorinnen arbeiteten lange im „WORD-Modus“
Anzahl BRAT-annotierter Dokumente	1022 (davon 52 5fach) + 15 kuratiert	550	850	

Überblick

- 1 Aufgabe(n)
- 2 Annotation
 - Begriffe
 - Korpus
 - Annotationsvorgang
- 3 Automatische Extraktion
- 4 Auswertung
 - Centroid
 - IAA & Systemperformanz

Automatische Extraktion

- Anpassen des JCoRE BIONLP *reader* und *writer* für die Medizin-Daten
- Erkennen von Medikamenten durch *Rote Liste Gazetteer* (JCoRE LINGPIPE)
- Erkennen der anderen Angaben (außer “Grund”) durch reguläre Ausdrücke
- Zuordnen der zusätzl. Angaben zu Medikamenten (einfache Fenster-basierte Heuristik)
- “Grund”-Erkennung nicht durchgeführt

Überblick

- 1 Aufgabe(n)
- 2 Annotation
 - Begriffe
 - Korpus
 - Annotationsvorgang
- 3 Automatische Extraktion
- 4 Auswertung
 - Centroid
 - IAA & Systemperformanz

Centroid

Beispiel:

- 5 Annotationsquellen

```
...-E-s-w-u-r-d-e-e-l-n-e-A-l-p-r-o-s-t-a-d-i-l-I-n-f-u-s-i-o-n-s-t-h-e-r-a-p-l-e-b-e-g-o-n-n-e-n-...
0 0 0 0 0 0 0 0 0 0 0 0 4 4 4 4 4 4 4 4 2 2 2 2 2 2 2 1 1 1 1 1 1 1 1 1 0 0 0 0 0 0 0 0
```

- *threshold* (t): die zu erreichende Anzahl an übereinstimmenden Kern-Annotationen
- *boundary* (b): Wert bei dem Verteilung um den Kern abgeschnitten wird (exklusiv)

F-Score – *One vs. All* für 52 Dokumente



- Inter Annotator Agreement (4 vs. 1):
 - ($t \leq 75\%$): 0.90 – 0.96; $\uparrow b \Rightarrow \downarrow \text{Precision}, \uparrow \text{Recall}$
 - ($t = 100\%$): 0.86 – 0.90; changing b doesn't change values
- Systemperformanz (JUMEX vs. 5):
 - 0.49 – 0.50

F-Score – *One vs. All* für 52 Dokumente

Dosierung

- Inter Annotator Agreement (4 vs. 1):
 - ($t \leq 75\%$): 0.91 – 0.98; $\uparrow b \Rightarrow \downarrow \text{Precision}, \uparrow \text{Recall}$
 - ($t = 100\%$): 0.89 – 0.94; changing b doesn't change values
- Systemperformanz (JUMEX vs. 5):
 - 0.81 – 0.83

F-Score – *One vs. All* für 52 Dokumente

Häufigkeit

- Inter Annotator Agreement (4 vs. 1):
 - ($t \leq 75\%$): 0.91 – 0.98; changing b doesn't change values
 - ($t = 100\%$): 0.90 – 0.96; changing b doesn't change values
- Systemperformanz (JUMEX vs. 5):
 - 0.81 – 0.83

F-Score – *One vs. All* für 52 Dokumente

Modus

- Inter Annotator Agreement (4 vs. 1):
 - ($t \leq 75\%$): 0.69 – 0.85; $\uparrow b \Rightarrow \downarrow \text{Precision}, \uparrow \text{Recall}$
 - ($t = 100\%$): 0.57 – 0.77; $\uparrow b \Rightarrow \downarrow \text{Precision}, \downarrow \text{Recall}$
- Systemperformanz (JUMEX vs. 5):
 - 0.19 – 0.22

F-Score – *One vs. All* für 52 Dokumente

Dauer

- Inter Annotator Agreement (4 vs. 1):
 - ($t \leq 75\%$): 0.66 – 0.78; $\uparrow b \Rightarrow \downarrow \text{Precision}, \uparrow \text{Recall}$
 - ($t = 100\%$): 0.58 – 0.74; $\uparrow b \Rightarrow \downarrow \text{Precision}, \downarrow \text{Recall}$
- Systemperformanz (JUMEX vs. 5):
 - 0.30 – 0.34

F-Score – *One vs. All* für 52 Dokumente

Grund

- Inter Annotator Agreement (4 vs. 1):
 - ($t \leq 75\%$): 0.27 – 0.69; $\uparrow b \Rightarrow \downarrow \text{Precision}, \rightarrow \text{Recall}$
 - ($t = 100\%$): 0.17 – 0.49; $\uparrow b \Rightarrow \downarrow \text{Precision}, \downarrow \text{Recall}$
- Systemperformanz (JUMEX vs. 5):
 - keine Erkennung

JULIELAB@SMITH

Annotation von Medikationsinformationen in deutschsprachigen Entlassbriefen

Franz Matthies

Lehrstuhl für Computerlinguistik
Institut für Germanistische Sprachwissenschaft
Friedrich-Schiller-Universität Jena
www.julielab.de

29. Juni 2017