

Block VI: Datenformate, Terminologien und Metadaten

Übung Terminologieannotierung mit Annotool

Matthias Löbe

Institut für Medizinische Informatik, Statistik und Epidemiologie (IMISE), Universität Leipzig

Ziel der Übung

- Praktische Erfahrung in der Annotierung von Datenelementen eines realen Datensatzes für ein medizinisches Forschungsvorhaben
- Annotierung = „Verschlagwortung“ mit einem oder ggf. mehreren Konzepten aus einer Terminologie mit einer Mapping-Relation (welche die Güte der Übereinstimmung angibt)
- Verwendung von Konzepten aus medizinischen Terminologien mit relevanter Verbreitung (LONC, SNOMED CT, ATC)
- Kritische Bewertung und Diskussion der Lösung
- Lernziel: Kennenlernen des Prozesses, seiner Stärken und Schwächen

Szenario: POLAR_MI

- POLAR_MI: POLypharmazie, Arzneimittelwechselwirkungen und Risiken
- POLAR_MI ist einer der übergreifenden Use Case der MII, d.h. Standorte aller vier Konsortien sind daran beteiligt
- Ziel des Use Cases: Detektion von Gesundheitsrisiken bei Polymedikation multimorbider Patienten
- Wir wollen die „Pilotabfrage 1.3“ (in ganz leicht abgewandelter Form) annotieren
- Die Polarabfragen sollen in 15 Universitätskliniken ausgeführt werden
 - Sie müssen 15 Hauskataloge unterstützen bzw. auf ein zentrales Modell „mappen“
 - Ein Schritt in diese Richtung ist die semantische Fundierung mit Konzepten aus Terminologien



Abzufragende Datenelemente Pilotabfrage 1.3

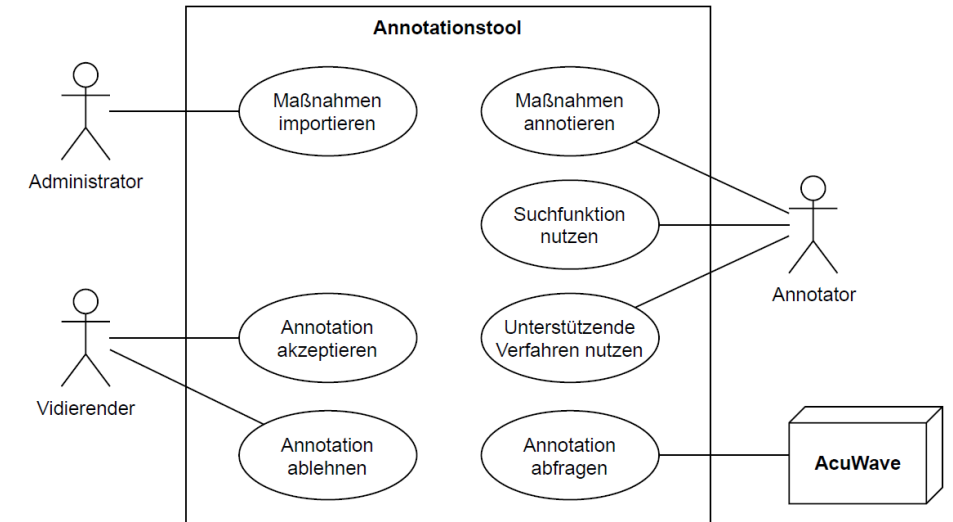
Variablenname	Beschreibung	Einheit
Alter	Alter bei Krankenhausaufnahme (abrunden auf Jahr)	Jahr
Geschlecht	[Geschlecht]	
Körpergewicht	[Gewicht]	kg
Körpertemperatur	[Temperatur]	°C
Systolischer Blutdruck	[SBP]	mmHg
GFR	Glomerular Filtration Rate (GFR) in ml/min	ml/min
CrCl	Creatinine Clearance (CrCl) in ml/min	ml/min
CrCl CG	Creatinine Clearance (CrCl) in ml/min nach Cockcroft&Gault	ml/min
GFR2	Glomerular Filtration Rate (GFR) in ml/min/1,73m2	ml/min/1,73m2
Zeitpunkt	Zeitpunkt der Medikationserhebung	
Metformin	Metformin	
Dronedaron	Dronedaron	
Dabigatran etexilat	Dabigatran etexilat	
Eplerenon	Eplerenon	
Ibuprofen	Ibuprofen	
Gabapentin	Gabapentin	
Allopurinol	Allopurinol	
Pregabalin	Pregabalin	
Sitagliptin	Sitagliptin	

Wiederholung der nutzbaren Terminologien

- Wir verwenden hier drei wichtige und einschlägige Terminologien
- LOINC: System zur Beschreibung von Laborwerten (siehe Folien 17-20)
- ATC: Klassifiziert pharmakologische Wirkstoffe (siehe Folie 22)
- SNOMED CT: Referenzterminologie mit einer sehr breiten Abdeckung der Medizin (siehe Folien 30-32)

„Annotool“

- Entwicklung des Datenintegrationszentrums Bonn
- Datenkataloge (leer) hochladen (Administratoren) >> Liste aller Datenelemente erscheint in Übersicht
- Annotatoren (Redakteure) verbinden die Datenelemente mit Konzepten aus den drei Terminologien
- Vidierende (Qualitätssicherer) geben die Annotationen frei
- Derzeit im Testbetrieb
- Sie sind ein Annotator!



Quelle: **Thomas Wiktorin**. Semantische Annotation im Gesundheitswesen. Prototypische Entwicklung und Evaluation eines kollaborativen Werkzeugs zur semantischen Annotation medizinischer Daten. Masterarbeit (2021).

Motivierendes Beispiel: GECCO Plus Forschungsdatensatz des NUM

GECCO PLUS

Anamnese / Risikofaktoren

Chronische Lungenerkrankungen

Herz-Kreislauf-Erkrankungen

Chronische Lebererkrankungen

Rheumatische/immunologische Erkrankungen

Bestehende HIV-Infektion

Organtransplantiert

Diabetes

Aktive Tumor-/Kreislauferkrankungen

Raucherstatus

Chronische neurologische oder psychiatrische Erkrankungen

Bestehende Sauerstoff- oder Beatmungstherapie

Chronische Nierenerkrankungen

Reiseaktivität

Magengeschwüre

Implantationen

DNR-Anordnung

Bildgebung

Demographie

Epidemiologische Faktoren

Chronische Lungenerkrankungen

Id

covid19f-dataelement-2.1130

Version

2020-08-12

Status

Entwurf

Versions-Label

Beschreibung

Leidet der/die Patient*in unter einer chronischen Lungenerkrankung? An welcher chronischen Lungenerkrankung leidet der/die Patient*in?

Beziehungen

Spezialisierung

Chronische Lungenerkrankungen (1.0) - covid19f-dataelement-1.1130 vom 2020-04-08

Kommentar

FHIR-Mapping: Condition

Terminologie Assoziation

Code	Wiedergabename	Codesystem
J98.4	Sonstige Veränderungen der Lunge	ICD-10 International (WHO) (German Variant)
413839001	chronische langziekte (aandoening)	SNOMED Clinical Terms

Wert

Typ

Kode

Value Set Assoziation

extensible

Chronische Lungenerkrankungen ICD (Dynamisch)

required

Certainty of Absence (Dynamisch)

required

Uncertainty of Presence (Dynamisch)

required

DataAbsentReason (Dynamisch)

required

Certainty of Presence (Dynamisch)

extensible

Chronische Lungenerkrankungen (Dynamisch)

Konzepte (Click the value set for full detail. Zeigt 20 von 73)

Konzept	Code	Wiedergabename	Codesystem
Asthma	J45.9 195967001	Asthma bronchiale, nicht näher bezeichnet Asthma (disorder)	ICD-10 International (WHO) (German Variant) SNOMED Clinical Terms
COPD	J44.9 13645005	Chronische obstruktive Lungenkrankheit, nicht näher bezeichnet Chronic obstructive lung disease (disorder)	ICD-10 International (WHO) (German Variant) SNOMED Clinical Terms
Lungenfibrose	J84.1 51615001	Sonstige interstitielle Lungenkrankheiten mit Fibrose Fibrosis of lung (disorder)	ICD-10 International (WHO) (German Variant) SNOMED Clinical Terms
Lungenhochdruck/pulmonale Hypertonie	I27.0 70995007	Primäre pulmonale Hypertonie Pulmonary hypertension (disorder)	ICD-10 International (WHO) (German Variant) SNOMED Clinical Terms
OHS	E66.29 190966007	Übermäßige Adipositas mit alveolärer Hypoventilation : Grad oder Ausmaß der Adipositas nicht näher bezeichnet Extreme obesity with alveolar hypoventilation (disorder)	ICD-10 International (WHO) (German Variant) SNOMED Clinical Terms
Schlafapnoe	G47.3 73430006	Schlafapnoe Sleep apnea (disorder)	ICD-10 International (WHO) (German Variant) SNOMED Clinical Terms

<https://art-decor.org/art-decor/decor-datasets--covid19f-?id=&effectiveDate=&conceptId=&conceptEffectiveDate=>

Folie 7 · Block VI: Datenformate, Terminologien und Metadaten – Übung Annotool · Matthias Löbe · 05.07.2022

TMF school 2022

Zugriff

- Webbasierte Applikation (mit Bugs), getestet mit Chrome, Edge, Firefox
- URL: <http://nisa3zjndxmyk33y.myfritz.net:8103/>
- Login: **hinterlegte E-Mail-Adresse** bei der Registrierung zur TMF-School
- Passwort: **demo** (für alle gleich)

Kurzanleitung Annotool: Katalog bereits hochgeladen

Semantische Annotation

Matt Demon

Logout

Annotation

Übersicht

+ Annotation

Übersicht über alle Annotationen am Standort IKF-Löbe2

Lokale Id	Typ	Bezeichnung	Einheit	Normwert	Abnahmeart	Dosis	Applikationsart	Letzte Änderung	Status	Details	Annotieren
[LOCAL ID]	[TYPE]	[DESCRIPTION]	[UNIT]	[STANDARD VALUE]	[DRAW TYPE]	[DOSE]	[ADMINISTRATI... FORM]	23.06.2021 11:33	In Bearbeitung	Details	Annotieren
Alter	Demographie	Alter bei Krankenhausauf... auf Jahr)	Jahr					23.06.2021 11:33	In Bearbeitung	Details	Annotieren
Geschlecht	Demographie	[Geschlecht]						23.06.2021 11:33	In Bearbeitung	Details	Annotieren
Körpergewicht	Demographie	[Gewicht]	kg					23.06.2021 11:33	In Bearbeitung	Details	Annotieren
Körpertemperat...	Demographie	[Temperatur]	°C					23.06.2021 11:33	In Bearbeitung	Details	Annotieren
Systolischer Blutdruck	Demographie	[SBP]	mmHg					23.06.2021 11:33	In Bearbeitung	Details	Annotieren
GFR	Laborwerte	Glomerular Filtration Rate (GFR) in ml/min	ml/min					23.06.2021 11:33	In Bearbeitung	Details	Annotieren
CrCl	Laborwerte	Creatinine Clearance (CrCl) in ml/min	ml/min					23.06.2021 11:33	In Bearbeitung	Details	Annotieren
CrCl_CG	Laborwerte	Creatinine Clearance (CrCl) in ml/min nach Cockcroft&Gault	ml/min					23.06.2021 11:33	In Bearbeitung	Details	Annotieren
GFR2	Laborwerte	Glomerular Filtration Rate (GFR) in ml/min/1,73m2	ml/min/1,73m2					23.06.2021 11:33	In Bearbeitung	Details	Annotieren
Zeitpunkt	Medikamentation	Zeitpunkt der Medikationserh...						23.06.2021 11:33	In Bearbeitung	Details	Annotieren
Metformin	Medikamentation	Metformin						23.06.2021 11:33	In Bearbeitung	Details	Annotieren
Dronedaron	Medikamentation	Dronedaron						23.06.2021 11:33	In Bearbeitung	Details	Annotieren

Kurzanleitung Annotool: neue Annotation

Semantische Annotation Matt Demon Logout

Annotation
Übersicht
+ Annotation

Lokale Id	Typ	Bezeichnung	Einheit	Bearbeitungsstand
Alter	Demographie	Alter bei Krankenhausaufnahme (abrunden auf Jahr)	Jahr	23.06.2021 11:33 Item auswählen

SNOMED CT
+

LOINC
+

ATC
+

Knowledge-Mappings Cross-Mappings Mapping zwischenspeichern Mapping abschicken

SNOMED CT Suche Hierarchie

Mindestens 3 Zeichen eingeben Alle Kategorien

SNOMED CT Konzept Mapped

Direkte Eingabe, wenn man einen Code schon kennt

Auswahl der Terminologie zur Suche nach Bezeichnern

Kurzanleitung Annotool: Suche

The screenshot shows the Annotool interface with the following components:

- Header:** Semantische Annotation, Matt Demon, Logout
- Left Sidebar:** Annotation, Übersicht, + Annotation
- Top Table:** Table with columns: Lokale Id, Typ, Bezeichnung, Einheit, Bearbeitungsstand. Row: Alter, Demographie, Alter bei Krankenhausaufnahme (abrunden auf Jahr), Jahr, 23.06.2021 11:33, Item auswählen
- Search Area:** Three boxes for SNOMED CT, LOINC, and ATC, each with a "+" button. Below them are buttons: Knowledge-Mappings, Cross-Mappings, Mapping zwischenspeichern, Mapping abschicken.
- Search Results:** A dropdown menu is open for LOINC. The search term "age" is entered. The results table is shown below:

LOINC-Code	Long Common Name	Mapped
30525-0	Age	0
54505-3	Language	0
46251-5	Age group	0
91887-0	Agency type	0
85151-9	XR Bone age	0
75500-9	Triage note	0
45972-7	Language Set	0
39016-1	Age at death	0
91698-1	Volunteer age	0
91691-6	Personnel Age	0

Red annotations highlight the "LOINC" dropdown and the "30525-0 Age" row in the results table. A red arrow points from the text "Aktuell nur englische Suche" to the search area. Another red arrow points from the text "Auswahl: Erweiterte Attribute werden angezeigt" to the selected row.

Details for the selected LOINC code 30525-0:

- Age**
- LOINC-Code:** 30525-0
- Component:** Age
- Property:** Time
- Time:** Pt
- System:** <https://loinc.org/30525-0/>
- Buttons:** Als Mapping hinzufügen

Kurzanleitung Annotool

Semantische AnnotationMatt DemonLogout

Annotation

Übersicht

+ Annotation

Lokale Id	Typ	Bezeichnung	Einheit
Alter	Demographie	Alter bei Krankenhausaufnahme (abrunden auf Jahr)	Jahr

SNOMED CT

+

LOINC

+

ATC

+

Bearbeitungsstand23.06.2021 11:33Item auswählen

Knowledge-MappingsCross-MappingsMapping zwischenspeichernMapping abschicken

SNOMED CTSuche

Hierarchie

age

Alle Kategorien

SNOMED CT Konzept	Mapped
Age (qualifier value)	0
Age (observable entity)	0
Cage (physical object)	0
Sage (substance)	0
Agent (metadata)	0
Agene (substance)	0
Pager (physical object)	0
Image (metadata)	0
Phage (organism)	0
Stage (metadata)	0

Age

SCTID: 397669002

Synonyme:

Age - descriptor

Age

Age (qualifier value)

Lokale Bezeichnung	Standort
Age	MDM-Portal
Gender Age	MDM-Portal
Adult Age	MDM-Portal
Alter mindestens 18 Jahre	MDM-Portal
age at least 18 Years	MDM-Portal
Alter	MDM-Portal
18 years of age or older	MDM-Portal
age > 18 years	MDM-Portal
age ≥ 18 years	MDM-Portal
gender, age	MDM-Portal

Als Mapping hinzufügen

2. Annotation speichern

1. Eine Annotation hinzufügen

TMF school 2022

Kurzanleitung Annotool

Semantische AnnotationMatt DemonLogout

Annotation

Übersicht

+ Annotation

Lokale Id	Typ	Bezeichnung	Einheit	Bearbeitungsstand	
Alter	Demographie	Alter bei Krankenhausaufnahme (abrunden auf Jahr)	Jahr	23.06.2021 12:31	Item auswählen

SNOMED CT

397669002

Disjoint

—

424144002

Equal

—

+

LOINC

30525-0

Equal

—

+

ATC

+

Knowledge-Mappings

Cross-Mappings

Mapping zwischenspeichern

Mapping abschicken

SNOMED CT

Suche

Hierarchie

age

Alle Kategorien

SNOMED CT Konzept	Mapped
Age (qualifier value)	0
Age (observable entity)	0
Cage (physical object)	0
Sage (substance)	0
Agent (metadata)	0
Agene (substance)	0
Pager (physical object)	0
Image (metadata)	0
Phage (organism)	0
Stage (metadata)	0

Current chronological age

SCTID: 424144002

Synonyme:

Current chronological age (observable entity)

Current chronological age

Age

Lokale Bezeichnung	Standort
Age	MDM-Portal
Gender Age	MDM-Portal
Adult Age	MDM-Portal
Alter mindestens 18 Jahre	MDM-Portal
age at least 18 Years	MDM-Portal
Alter	MDM-Portal
18 years of age or older	MDM-Portal
age > 18 years	MDM-Portal
age ≥ 18 years	MDM-Portal
gender, age	MDM-Portal

Als Mapping hinzufügen

Annotation beenden und
zur Qualitätssicherung vorlegen
(dann nicht mehr bearbeitbar)

Folie 14 · Block VI: Datenformate, Terminologien und Metadaten – Übung Annotool · Matthias Löbe · 05.07.2022

TMF school 2022

Kurzanleitung Annotool: Fortgeschrittene Funktionen 1

Semantische Annotation

Matt Demon

Logout

Annotation

Übersicht

+ Annotation

Lokale Id	Typ	Bezeichnung	Einheit
Alter	Demographie	Alter bei Krankenhausaufnahme (abrunden auf Jahr)	Jahr

SNOMED CT

397669002

Disjoint

424144002

Equal

+

LOINC

30525-0

Equal

+

ATC

+

Knowledge-Mappings

Cross-Mappings

Mapping zwischenspeichern

Mapping abschicken

SNOMED CT

Suche

Hierarchie

age

Alle Kategorien

SNOMED CT Konzept	Mapped
Age (qualifier value)	0
Age (observable entity)	0
Cage (physical object)	0
Sage (substance)	0
Agent (metadata)	0
Agene (substance)	0
Pager (physical object)	0
Image (metadata)	0
Phage (organism)	0
Stage (metadata)	0

Current chronological age

SCTID: 424144002

Synonyme:

Current chronological age (observable entity)

Current chronological age

Age

Als Mapping hinzufügen

Lokale Bezeichnung	Standort
Age	MDM-Portal
Gender Age	MDM-Portal
Adult Age	MDM-Portal
Alter mindestens 18 Jahre	MDM-Portal
age at least 18 Years	MDM-Portal
Alter	MDM-Portal
18 years of age or older	MDM-Portal
age > 18 years	MDM-Portal
age ≥ 18 years	MDM-Portal
gender, age	MDM-Portal

Suchunterstützung durch Anzeige existierender Mappings

Fortgeschrittene Funktionen 2: Knowledge-Mappings

Knowledge-Mappings bieten dem Annotator an, die Annotationen ähnlicher bereits annotierter Maßnahmen zu übernehmen

Knowledge-Mappings für

Lokale Id	Typ	Bezeichnung	Einheit	Normwert	Abnahmeart	Dosis	Applikationsart
Geschlecht	Demographie	[Geschlecht]	null	null	null	null	null
Gewichtung	Gewichtung	Gewichtung	Gewichtung	Gewichtung	Gewichtung	Gewichtung	Gewichtung
- 0 +	- 1 +	- 1 +	- 0 +	- 0 +	- 0 +	- 0 +	- 0 +

Maximale Levenshtein-Distanz
- 100 + [Mappings neu berechnen](#)

Ähnliche Annotationsitems

Lokale Id	Typ	Bezeichnung	Einheit	Normwert	Abnahmeart	Dosis	Applikationsart	Standort	Levenshtein
Geschlecht	Demographie	[Geschlecht]						IKF-Hempel	0
Körpergewicht	Demographie	[Gewicht]	kg					IKF-Haendly	5

[Mappings](#)
[Mappings](#)

Mappings ansehen und sinnvolle in eigene Annotation übernehmen

Anzeige „ähnlicher“ Datenelemente anhand der Levenshtein-Distanz der Bezeichner (profitiert von einer vollen Datenbank und „Schwarmintelligenz“)

[Schließen](#)

Fortgeschrittene Funktionen 3: Crossmappings

Existiert bereits eine Annotation (hier SNOMED CT), können über eine UMLS-Verknüpfung automatisch „richtige“ korrespondierende Konzepte gefunden werden (hier in ATC)

The screenshot shows the 'Semantische Annotation' interface with a modal dialog titled 'Mögliche Cross-Mappings des UMLS-Metathesaurus'. The dialog has tabs for 'ATC', 'LOINC', and 'SNOMED CT'. The 'ATC' tab is active, showing 'Sitagliptin' with ATC-Code 'A10BH01' and 'Defined Daily Doses' of '0,1 g O'. A red box highlights a table of existing mappings in the bottom right corner, with a red arrow pointing to it from the text 'Suchunterstützung durch Anzeige existierender Mappings'.

Lokale Bezeichnung	Standort
Sitagliptin	MDM-Portal
Thiazolidinediones exenatide sitagliptin Pramlintide	MDM-Portal
receiving tzd's, exenatide, sitagliptin or pramlintide therapy.	MDM-Portal
Diet therapy Exercise Sitagliptin Dose Stable Oral hypoglycemic Hemoglobin A1c measurement	MDM-Portal
currently being treated with diet/exercise and on stable dose of sitagliptin for ≥ 4 weeks, on any additional oral hypoglycemic agent (aha) in the prior 8 weeks, and hba1c ≥ 6.0% and ≤ 9.0%; or	MDM-Portal

Aufgabe

Annotieren Sie alle Konzepte aus der POLAR-Pilotanfrage 1.3!

- Es gibt noch keine Musterlösung 😊
- Es kann 0 bis n Treffer in jeder Terminologie geben.
- Falls es mehr als einen Treffer gibt, soll der oder die „besten“ Treffer ausgesucht werden.
 - Wenn bspw. einfach nur das Medikament „Aspirin“ (Acetylsalicylsäure) annotiert werden soll, kann es als Schmerzstiller und als Entzündungshemmer in Terminologien gelistet sein, aber eine Aspirinallergie wäre eine falsche Annotation
- Notieren Sie kurz ihre Beobachtungen für die anschließende Diskussion!
- (Letzte Bemerkung: Das Szenario ist nicht unrealistisch. Es gibt nur sehr wenige Experten, daher wird mit Einsteigern vorannotiert.)

Diskussion der Lösung und Kritik