

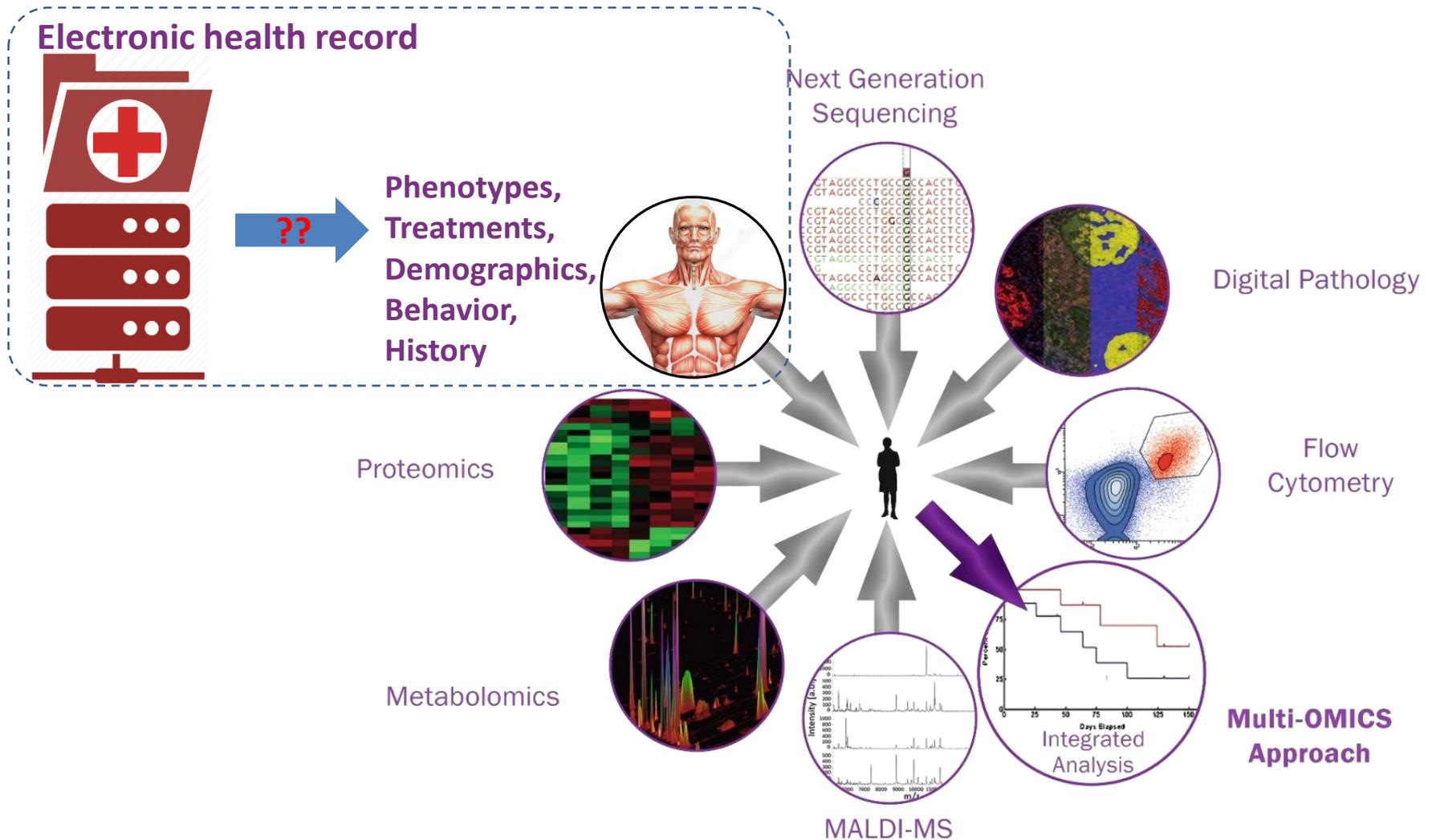
"Pat. mit rez. HWI u. VUR" -
Herausforderungen medizinischer
Sprache für Datenintegration in der
Systemmedizin

Stefan Schulz

Medizinische Universität Graz

stefan.schulz@medunigraz.at

Biomarkerforschung: Systemmedizinischer Ansatz



Natürliche Sprache in der Medizin

Natürliche Sprache in der Medizin

- Wichtigster Träger klinischer Information, optimiert auf menschliche Kommunikation
- Notwendigkeit strukturierter und kodierter klinischer Daten
- Computer müssen klinische Sprache "verstehen"
- NLP (Natural Language Processing):
 - Finden relevanter Dokumente (Text Retrieval)
 - Informationsextraktion aus Dokumenten
- Zahlreiche Herausforderungen der Kliniksprache...

knapp, potentiell mehrdeutig, ...

Patient?
Pathologie?

rezent?
rezidivierend?

Pat. mit rez.
HWI und
VUR

Harnwegsinfekt ?
Hinterwandinfarkt?

Vesicoureteral
reflux



variantenreich...

Colon-Ca

Kolon-Ca

Kolon-
karzinom

Colon-
carcinom

Colon-
Karzinom

Kolonkrebs

Dickdarm-
krebs

Dickdarm-Ca

Malignom des
Kolon

Dickdarm-
karzinom

Dickdarm-Ca

Bösartige
Neubildung
am Dickdarm

Bösartiger
Dickdarm-
tumor

maligne
Neoplasie des
Dickdarms

Karzinom des
Dickdarms

maligne NPL
des Colon

fehlertolerant...

Simvastatin

Sinvastatin

Simvastastin

Simvastain

Simvastad

Simbastatin

Simavstatin

Simavastatin

Simastatin

Symvastatin

Simvastation

Simvaststin

Simvatatin

Simvatin

Simvatstain

Simvstatin

wenig qualitätsgesichert

- **Mondscheinhaut** gut durchblutet
- die Patientin zeigte eine **geistliche** retardierung
- im Verlauf Kontrolle durch **Bett-CT**
- Pat. wurde zunehmend **in kompleient**
- heraushängen der **lunge** nach rechts
- Zustand nach **pleuraler Prostataresektion**
- Diagnostik bei **parasitierender** Leukozytose
- Diagnose: **mexikanische** Aortenklappe
- Diagnose: **Fischgerät** im Hals
- V.a. **Kreuz fährt** Jakob Krankheit...
- Untersuchungsbefund: Kopf/Hals: **großer Zähne**
- Cochlea-**Reh**-Implantation
- vollständig **alkoholisiertes** Hüftgelenk rechts
- Gabe eines **Wetterblockers**

"Googlen" in medizinischen Texten?

Makroskopie: "Resektat nach Whipple": Ein noch nicht eröffnetes Resektat, bestehend aus einem distalen Magen mit einer kleinen Kurvaturlänge von 9,5 cm und einer großen Kurvaturlänge von 13,5 cm, sowie einem duodenalen Anteil von 14 cm Länge. 2 cm aboral des Pylorus zeigt die Dünndarmwandung eine sanduhrartige Stenose. Im Magen- und Duodenallumen reichlich zähflüssiger Schleim, sanguinolent; die Schleimhaut ist insgesamt livide. Auf lamellierenden Schnitten zähfestes weißliches, teilweise nodulär konfiguriertes Gewebe, ohne das Gallengänge manifest werden. Der distale Anteil des Ductus pancreaticus ist leicht erweitert und von der Papilla Vateri aus 4,5 cm weit sondierbar, wobei er hier in einer peripankreatischen Narbenzone abbricht. Eine Gallengangsmündung läßt sich makroskopisch nicht abgrenzen. Die berichtete Duodenumstenose liegt 2,5 cm oral der Papilla Vateri und steht mit der beschriebenen Narbenzone in direktem Zusammenhang.

"Googlen" in medizinischen Texten?

Makroskopie: "Resektat nach Whipple": Ein noch nicht eröffnetes Resektat, bestehend aus einem distalen Magen mit einer kleinen Kurvaturlänge von 9,5 cm und einer großen Kurvaturlänge von 13,5 cm, sowie einem duodenalen Anteil von 14 cm Länge. 2 cm aboral des Pylorus zeigt die Dünndarmwandung eine sanduhrartige Stenose. Im Magen- und Duodenallumen reichlich zähflüssiger Schleim, sanguinolent; die Schleimhaut ist insgesamt livide. Auf lamellierenden Schnitten zähfestes weißliches, teilweise nodulär konfiguriertes Gewebe, ohne das Gallengänge manifest werden. Der distale Anteil des Ductus pancreaticus ist leicht erweitert und von der Papilla Vateri aus 4,5 cm weit sondierbar, wobei er hier in einer peripankreatischen Narbenzone abbricht. Eine Gallengangsmündung läßt sich makroskopisch nicht abgrenzen. Die berichtete Duodenumstenose liegt 2,5 cm oral der Papilla Vateri und steht mit der beschriebenen Narbenzone in direktem Zusammenhang.

**Dokument wird
gefunden mit den
Suchwörtern:
"Whipple", "Magen",
"Pylorus"**

"Googlen" in medizinischen Texten?

Makroskopie: "**Resektat** nach **Whipple**": Ein noch nicht eröffnetes **Resektat**, bestehend aus einem distalen **Magen** mit einer kleinen Kurvaturlänge von 9,5 cm und einer großen Kurvaturlänge von 13,5 cm, sowie einem duodenalen Anteil von 14 cm Länge. 2 cm aboral des **Pylorus** zeigt die Dünndarmwandung eine sanduhrartige Stenose. Im Magen- und **Duodenal**lumen reichlich zähflüssiger Schleim, sanguinolent; die Schleimhaut ist insgesamt livide. Auf lamellierenden Schnitten zähfestes weißliches, teilweise nodulär konfiguriertes Gewebe, ohne dass Gallengänge manifest werden. Der distale Anteil des **Ductus pancreaticus** ist leicht erweitert und von der Papilla Vateri aus 4,5 cm weit sondierbar, wobei er hier in einer **peripankreatischen** Narbenzone abbricht. Eine **Gallengangs**mündung lässt sich makroskopisch nicht abgrenzen. Die berichtete Duodenumstenose liegt 2,5 cm oral der Papilla Vateri und steht mit der beschriebenen Narbenzone in direktem Zusammenhang.

Dokument wird
gefunden mit den
Suchwörtern:
"Whipple", "Magen",
"Pylorus"

Keine Treffer für:
"Pankreatikoduodenektomie", "Resektion",
"Duodenum", "Zwölffingerdarm", "Pankreas",
"Bauchspeicheldrüse", "Gallengang", "Pankreasgang",
"Ductus pancreaticus", "Papille", "Magenresektion"

Ziel: Abbildung auf standardisierte Bedeutung

**St. p. TE eines
exulc.
sek.knot.SSM li US
dors. 5/11 Level IV
2,4 mm
Tumordurchm.
Sentinell LK ing.
li. tumorfr.**



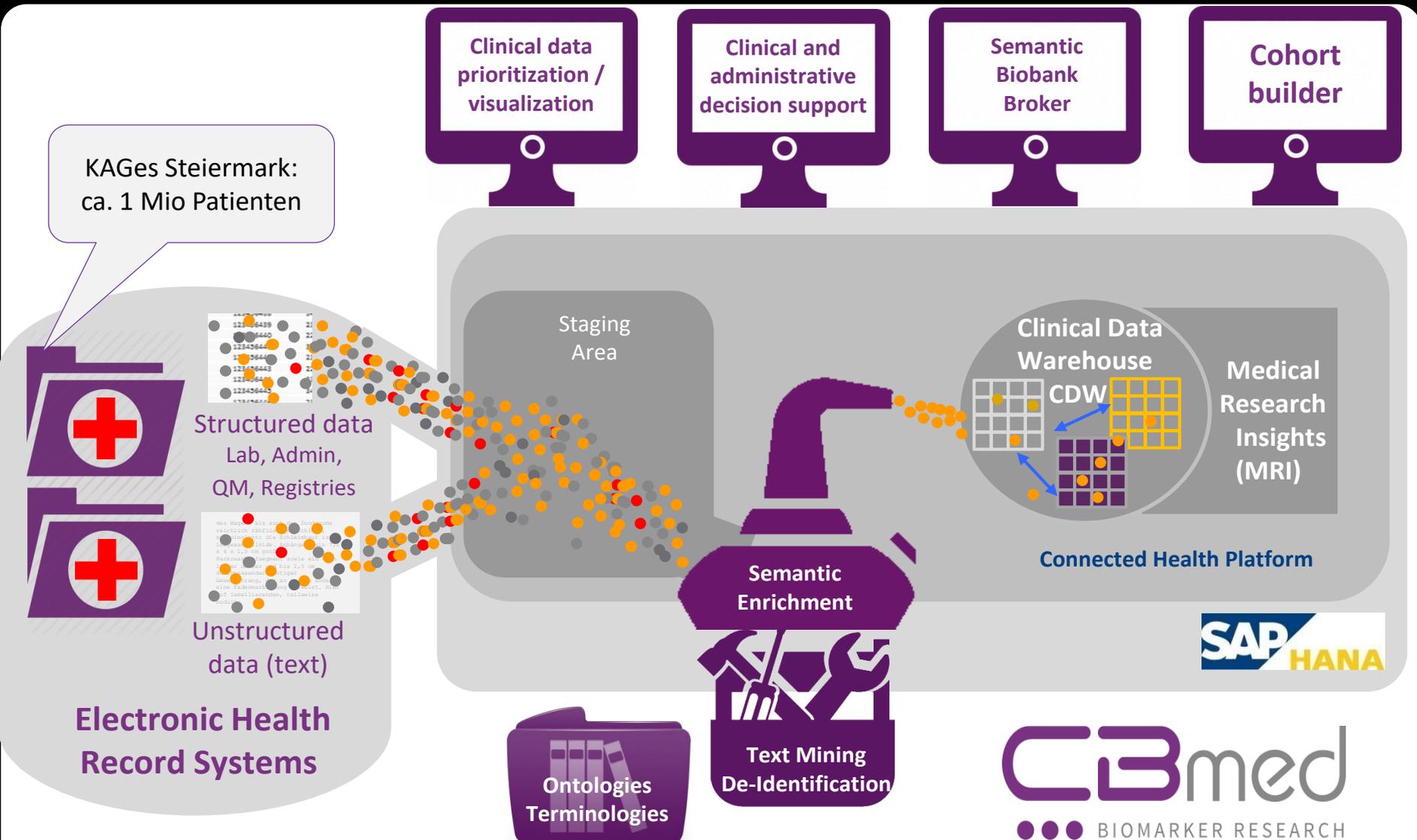
Ziel: Abbildung auf standardisierte Bedeutung

St. p. TE eines
exulc.
sek.knot.SSM li US
dors. 5/11 Level IV
2,4 mm
Tumordurchm.
Sentinell LK ing.
li. tumorfr.



Code (SNOMED CT, LOINC)	Wert	Kontext
254730000 Superficial spreading malignant melanoma of skin		392521001 History of
301889008 Excision of malignant skin tumor		392521001 History of
47224004 Skin of posterior surface of lower leg		
7771000 Left		
81827009 Diameter	2.41	
258673006 millimeter		
258403002 Lymph node level IV		
94339008 Secondary malignant neoplasm of inguinal lymph nodes		15240007 Current 2667000 Absent

Implementierung semantischer Technologien in CBmed – Biomarker Research, Graz



Ressourcen und Werkzeuge zur maschinellen Analyse von medizinischen Texten

Ressourcen

Ressourcen

- Lexikalisch-ontologische Ressourcen
 - Ontologien / Referenzterminologien
 - Interfaceterminologien
 - Klassifikationssysteme
- Informationsmodelle
- Korpora
 - annotierte Korpora: "supervised" Learning
 - nichtannotierte Korpora: "unsupervised" Learning

Ressourcen

- Lexikalisch-ontologische Ressourcen
 - **Ontologien / Referenzterminologien**
 - **Interfaceterminologien**
 - Klassifikationssysteme
- Informationsmodelle
- Korpora
 - **annotierte Korpora: "supervised" Learning**
 - **nichtannotierte Korpora: "unsupervised" Learning**

Beispiel: Terme in Kardiologie-Arztbriefen

Vorzugsterm (ICD, OPS)	Anzahl	Synonym	Anzahl
Aortenklappenstenose	3749	Aortenstenose	3126
Hirninfarkt	7	Schlaganfall	65
Elektrokardiogramm	0	EKG	12208
Koronare Herzerkrankung	331	KHK	18455
Nicht-ST-Hebungsinfarkt	498	NSTEMI	3839
Magnetresonanztomographie	2	NMR	17

Interface-terminologie

beschreibt
Sprache einer
Domäne

beschreibt die
Gegenstände
einer Domäne

Interface-terminologie

"Lungenkrebs"
"Bronchialkarzinom"

"Ca"
"Kalzium"
"Calcium"

"Ca"
"Krebs"
"Karzinom"

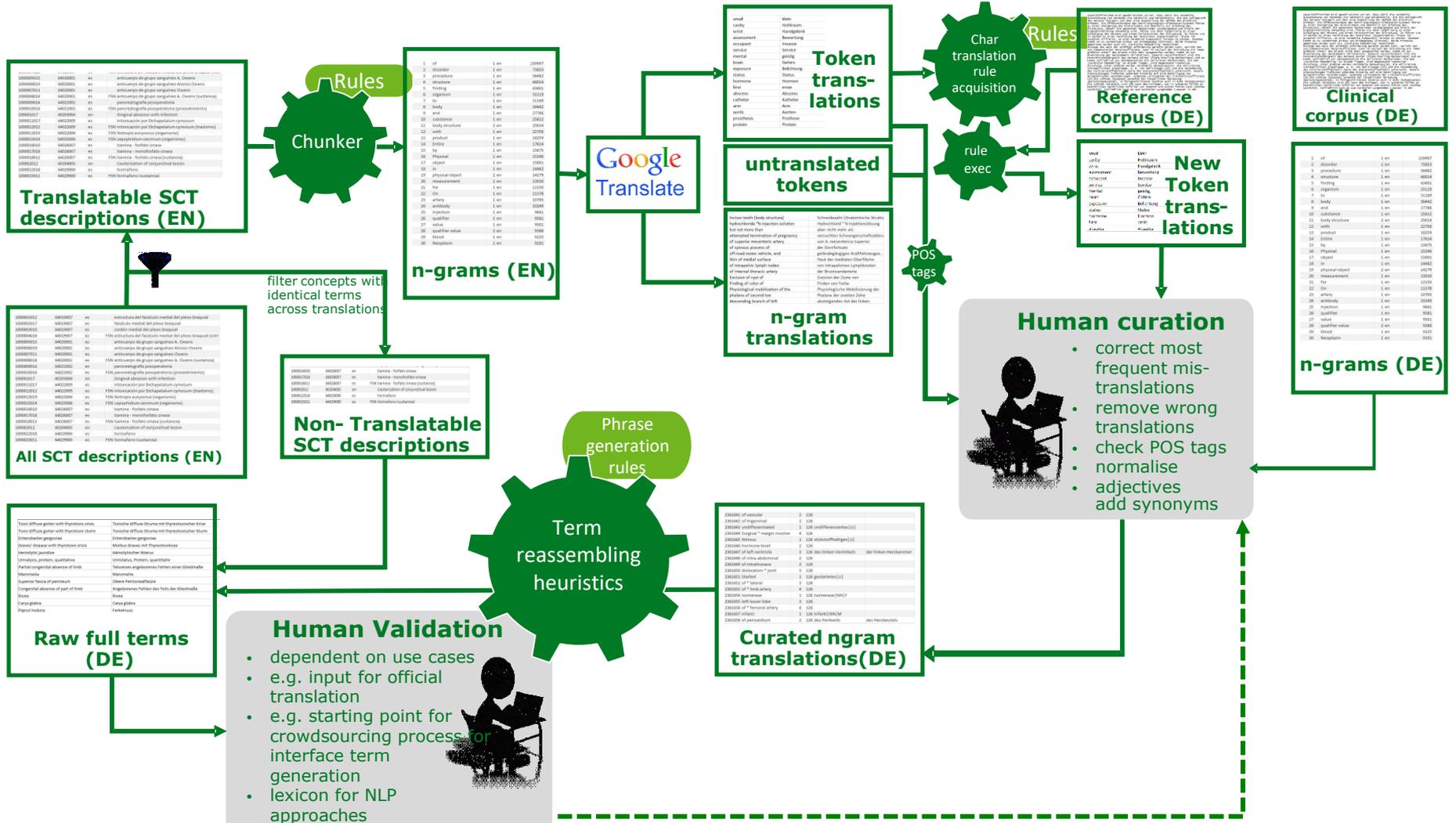
Ontologie

93880001
Primary malignant neoplasm of lung

5540006
Calcium (substance)

68453008 Carcinoma
(morph. abnormality)

MUG-GIT: Erstellung einer deutschen Interface-terminologie für SNOMED CT (II)



Automatische generierte Interfaceterminologie

20170315_240011_002	126952004	Neoplasm of brain	Gehirneubildung
20170315_240011_003	126952004	Neoplasm of brain	Neubildung des Hirns
20170315_240011_004	126952004	Neoplasm of brain	Hirnneubildung
20170315_240011_005	126952004	Neoplasm of brain	Neoplasie des Gehirns
20170315_240011_006	126952004	Neoplasm of brain	Gehirneoplasie
20170315_240011_007	126952004	Neoplasm of brain	Neoplasie des Hirns
20170315_240011_008	126952004	Neoplasm of brain	Hirneoplasie
20170315_240011_009	126952004	Neoplasm of brain	Neoplasma des Gehirns
20170315_240011_010	126952004	Neoplasm of brain	Gehirneoplasma
20170315_240011_011	126952004	Neoplasm of brain	Neoplasma des Hirns
20170315_240011_012	126952004	Neoplasm of brain	Hirneoplasma
20170315_241010_001	126953009	Neoplasm of cerebrum	Neubildung des Großhirns
20170315_241010_002	126953009	Neoplasm of cerebrum	Neoplasie des Großhirns
20170315_241010_003	126953009	Neoplasm of cerebrum	Neoplasma des Großhirns
20170315_242015_001	126954003	Neoplasm of frontal lobe	Neubildung des Frontallappens
20170315_242015_002	126954003	Neoplasm of frontal lobe	Neubildung des Lobus frontalis
20170315_242015_003	126954003	Neoplasm of frontal lobe	Neoplasie des Frontallappens
20170315_242015_004	126954003	Neoplasm of frontal lobe	Neoplasie des Lobus frontalis
20170315_242015_005	126954003	Neoplasm of frontal lobe	Neoplasma des Frontallappens
20170315_242015_006	126954003	Neoplasm of frontal lobe	Neoplasma des Lobus frontalis
20170315_243013_001	126955002	Neoplasm of temporal lobe	Neubildung des Temporallappens
20170315_243013_002	126955002	Neoplasm of temporal lobe	Neubildung des Lobus temporalis
20170315_243013_003	126955002	Neoplasm of temporal lobe	Neoplasie des Temporallappens
20170315_243013_004	126955002	Neoplasm of temporal lobe	Neoplasie des Lobus temporalis
20170315_243013_005	126955002	Neoplasm of temporal lobe	Neoplasma des Temporallappens

Ressourcen

- Lexikalisch-ontologische Ressourcen
 - **Ontologien / Referenzterminologien**
 - **Interfaceterminologien**
 - Klassifikationssysteme
- Informationsmodelle
- Korpora
 - **annotierte Korpora: "supervised" Learning**
 - **nichtannotierte Korpora: "unsupervised" Learning**

Beispiel: Annotierter Korpus (Entlassmedikation in Arztbrief)

Thrombo	Ass	100mg	0-1-0				
DrugName	DrugName	Strength	Regimen				
Sortis	80mg	0-0-1					
DrugName	Strength	Regimen					
Pantoloc	ret.	47,5mg	1	-	0	-	1
DrugName	DrugName	Strength	Regimen	Regimen	Regimen	Regimen	Regimen
Seloken	DA	2-0-0					
DrugName	DrugName	Regimen					
Oleovit	D3	1x	wöchentlich	(Do)			
DrugName	DrugName	Regimen	Regimen	Regimen			
Thyrex	10 mg	mg	1-0-0				
DrugName	Strength	Strength	Regimen				
Torasemid	1-0-1						
DrugSubstance	Regimen						
Antiflat	5mg	3	x	3	ml	bei	Blähungen
DrugName	Strength	Regimen	Regimen	Regimen	Regimen	Other	Other
Xatral	forte	0-0-0-1					
DrugName	DrugName	Regimen					
Dominal	25mg	1-0-0					
DrugName	Strength	Regimen					
Marcoumar	laut	Pass	bitte	um	Gerinnungskontrolle	beim	HA
DrugName	Other	Other	Other	Other	Other	Other	Other
Laevolac	3x2EL						
DrugName	Regimen						
Hypren	Kapseln	3x1					
DrugName	DoseForm	Regimen					
Bioflorin	500mg	p.	o.	1-0-0			
DrugName	Strength	Route	Route	Regimen			
Tavanic	40mg	s.c.	1x1	abends			
DrugName	Strength	Route	Regimen	Regimen			
Urosin	300	mg	dzt.	pausiert			
DrugName	Strength	Strength	Other	Other			

Beispiel: Nichtannotierter Korpus: N-Gramm-Modelle aus 30k Arztbriefen

- 9 → I00.0 · Kardiomyopathie · dilatativ
- 9 → Deflamat · b
- 9 → besteht · eine · vorbekannte
- 8 → keine · umschriebenen · Infiltrate.
- 8 → Fortimel
- 8 → keine · SVT · oder
- 8 → T-ASS · 0 · x · 000 · mg
- 8 → für · 0 · Wochen · empfohlen.
- 8 → Nicolan · 00mg · 0
- 8 → der · unteren · Atemwege
- 8 → zeigte · sich · eine · grenzwertig
- 8 → Zeitpunkt · der · Aufnahme · an
- 8 → bei · rezid.
- 8 → mit · NSTEMI · stationär
- 8 → erklärt · werden · kann.
- 8 → zur · geplanten · Simdax-Therapie
- 8 → Umbilicalhernie

Ressourcen und Werkzeuge zur maschinellen Analyse von medizinischen Texten

Werkzeuge

- NLP-Pipeline:
Erkennung von



- Dokumentenabschnitten
- Sätzen
- Phrasen
- Wortarten
- Fachtermini (einschl. Abkürzungen, Fehlern)
- Semantischen Relationen
- Kontexte sprachlicher Ausdrücke

Beispiel: Erkennen von Abkürzungen

Beispiel: Erkennen von Abkürzungen

A.

A.

A.

A.

A.

A.

A.

A.

Abkürzung oder Satzende?

einer Infektion mit Hepatitis A. Hierbei sind keine weiteren
ausgeprägte Stenose der A. Subclavia, die eine Indikation
auf Streptokokken vom Typ A. Nicht ausgeschlossen ist, dass
akuter Mangel an Vitamin A. Wir empfehlen Substitution
Symptomatik haben wir 1 A. Digimerck verabreicht, um die
nach Verlegung auf Station 6 A. Wir bedauern, ihnen keine
ein Adenokarzinom o. n. A. diagnostiziert worden. Dabei
Ausschluss von Hämophilie A. Die Koagulopathie konnte bis

Abkürzung oder Satzende?

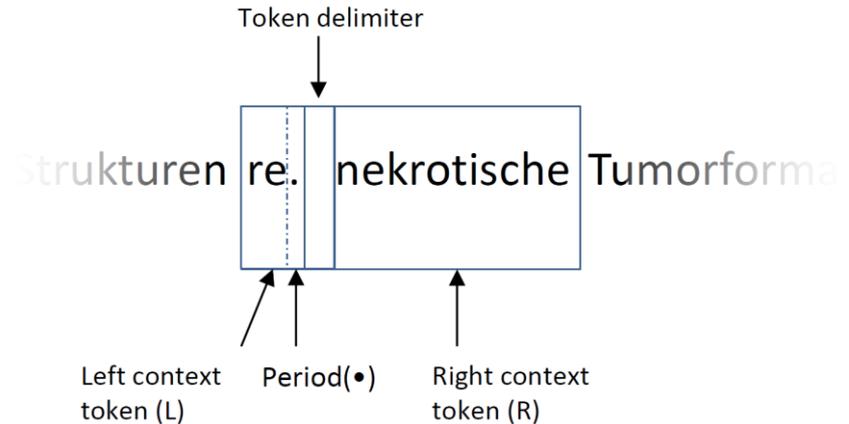
einer Infektion mit **Hepatitis A**. Hierbei sind keine weiteren
ausgeprägte Stenose der **A. Subclavia**, die eine Indikation
auf Streptokokken vom **Typ A**. Nicht ausgeschlossen ist, dass
akuter Mangel an **Vitamin A**. Wir empfehlen Substitution
Symptomatik haben wir **1 A. Digimerck** verabreicht, um die
nach Verlegung auf Station **6 A**. Wir bedauern, ihnen keine
ein Adenokarzinom **o. n. A.** diagnostiziert worden. Dabei
Ausschluss von **Hämophilie A**. Die Koagulopathie konnte bis

Abkürzung oder Satzende?

- Klassifikationsproblem
 - Punkt ist Teil des linken Wortes, kein Satzende
 - Punkt ist Teil des linken Wortes, Satzende
 - Punkt ist nicht Teil des linken Wortes → Satzende

- Featureextraktion

- Wortlänge
- Wortklasse
- Korpus
- Regeln
- Dictionary
- Verteilung



Abkürzungserkennung

- Klassifikation mittels SVN
- F1-Werte nach Methode (kumulativ)

	Baseline	Regeln	Statistik	Scaling	Corpus	Länge	Wortklasse
Training	0.62	0.60	0.71	0.86	0.88	0.95	0.97
Test	0.60	0.60	0.71	0.83	0.96	0.93	0.95

Beispiel: Auflösen von Abkürzungen

Beispiel: Auflösen von Abkürzungen

- "dilat. Kardiomyopathie, hochgr. red. EF"

Beispiel: Auflösen von Abkürzungen

- "dilat. Kardiomyopathie, hochgr. red. EF"
- Wort N-gram Modell (aus 30000 Arztbriefen)

1035	dilat. Kardiomyopathie
1442	dilatative Kardiomyopathie
7	hochgr. red. EF
4	hochgradig reduzierte EF

Beispiel: Auflösen von Abkürzungen

- "dilat. Kardiomyopathie, hochgr. red. EF"
- Wort-N-gram Modell (aus 30000 Arztbriefen)

```
1035      dilat. Kardiomyopathie
1442      dilatative Kardiomyopathie
```

```
7         hochgr. red. EF
4         hochgradig reduzierte EF
```

- Web mining 

[Ejektionsfraktion – Wikipedia](#)

<https://de.wikipedia.org/wiki/Ejektionsfraktion> ▼ [Translate this page](#)

Die **Ejektionsfraktion (EF)** oder Auswurffraktion (auch Austreibungsfraktion) ist ein Maß für die ... 30 %
hochgradig eingeschränkt ... Eine **reduzierte** Ejektionsfraktion wird als objektivierbarer Parameter

Beispiel: Auflösen von Abkürzungen

- "Pat. mit rez. HWI und VUR"
- Wort-N-gram Modell

```
381 Pat. mit  
120 Patient mit  
  2 rez.  
707 rezenten  
468 rezidivierende
```

- Web mining 

Der **vesikoureterale Reflux (VUR)** - KidsDoc.at

www.kidsdoc.at/vesikoureteraler_reflux_vur.html ▼ Translate this page

Der **vesikoureterale Reflux (VUR)** ist ein Zurückfließen des Urins aus der Harnblase in den oberen Harntrakt (Harnleiter und Nierenbecken). Der Harn soll von ...

[PDF] **HWI Ursachen und Folgen Leoben 2012**

www.paediatric.at/.../HWI_Ursachen_Folgen_Leoben_2012.pdf ▼ Translate this page

Nov 30, 2012 - Harnwegsinfekt (HWI) und Harnwegsfehlbildungen ... Erbrechen und Durchfall seit heute ... **Vesico-ureteraler Reflux (VUR)**. • Obstruktive ...

[PDF] **Diagnose und Behandlung von ... - Swiss Paediatrics**

www.swiss-paediatrics.org/sites/default/files/10-13_0.pdf ▼ Translate this page

tik, Behandlung, Abklärung und Nachkontrol- ... nen, Säuglingen, Kindern und Jugendlichen bis. 16 Jahre. wegsinfektion noch das Vorliegen eines **VUR**.

[PDF] **Rezidivierender **Harnwegsinfekt** beim Kind - Krause und Pachernegg**

www.kup.at/kup/pdf/7439.pdf ▼ Translate this page

by J Oswald

bestätigten **HWI** um das 6. Lebensjahr liegt bei 7 % für Mädchen und 2 % bei Knaben. Die meisten ... der mit rezidivierenden **HWIs** ein **VUR** dokumentieren.

[PDF] **Harnwegsinfektionen im frühen Kindesalter - CME-Portal - MGO ...**

https://cme.mgo-fachverlage.de/uploads/exam/exam_94.pdf ▼ Translate this page

der ersten **HWI** deutlich jünger als die betroffenen Mädchen. Bei 36 % der Mädchen und bei 24 % der Jungen fand sich ein **vesikoureteraler Reflux (VUR)** ...

[PDF] **Harnwegsinfektion - Klinik und Poliklinik für Kinder- und ...**

kik.uniklinikum-leipzig.de/red_tools/dl_document.php?id=251 ▼ Translate this page

monas und Enterokokken je 2%. CAVE bei Jungs 30% Proteus ab 2. Lj. Prädisponierende Faktoren für Auftreten einer **HWI**: **Vesikoureteraler Reflux (VUR)**: 30% ...

Fazit

- Deutschsprachige lexikalisch / ontologische Ressourcen:
 - Nur teilweise Mappings zu internationalen Standards
 - SNOMED CT bisher nicht eingeführt
 - Gute Interface-Terminologie nur für ICD-10
- Klinische Korpora
 - Im Gegensatz zu USA keine öffentlich verfügbaren annotierten Medizinkorpora
- Tools:
 - NLP-Tools für deutsche Sprache nur begrenzt nutzbar für klinische Texte
- Firmen, spezialisiert auf deutschsprachige Kliniktexte:
 - Averbis GmbH
 - ID Berlin GmbH & Co. KGaA

Fragen?

Stefan Schulz

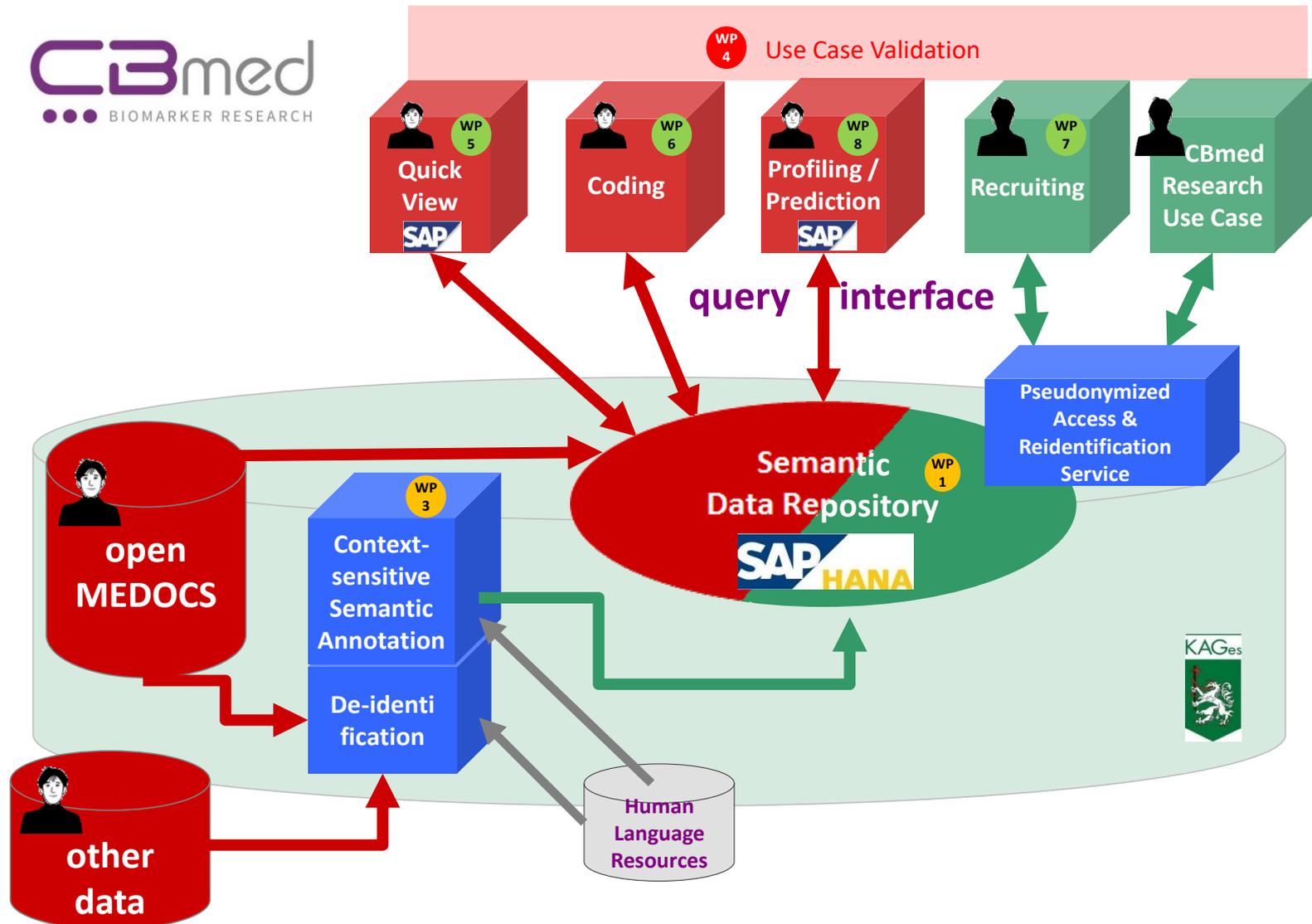
Medizinische Universität Graz

stefan.schulz@medunigraz.at

Ontologie und Informationskontext

Implementierung semantischer Technologien in Cbmed – Biomarker Research, Graz

Implementierung semantischer Technologien in Cbmed – Biomarker Research, Graz



Ressourcen für maschinelle Analyse geschriebener Sprache

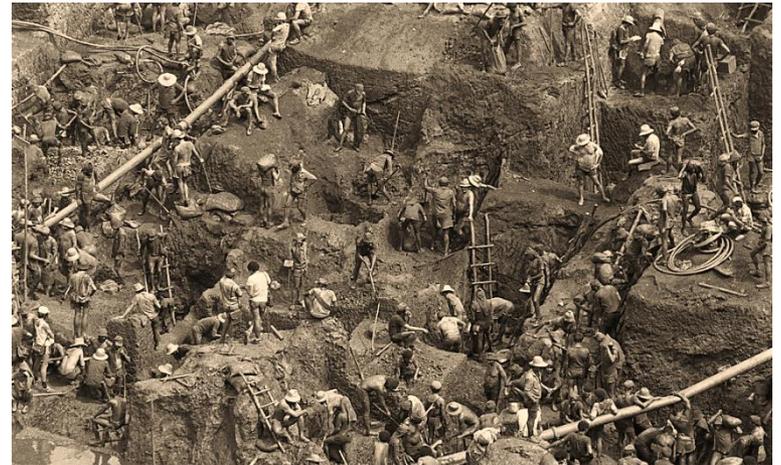
- Terminologien / Lexika
 - Fachtermini und deren Beziehungen, z.B. Synonyme:
"SSMM" = "Superficially spreading malignant melanoma"
- Ontologien
 - Axiomatische Beschreibungen, z.B.
MitralValve subclassOf partOf some Heart
- Grammatik / Regelwerke
 - z.B. Wortbildungsregeln: "-itis" = Entzündung, "-ektomie": chirurgische Entfernung
- Textkorpora (z.B. Kliniktexte, Veröffentlichungen, Web)
 - Annotierte Korpora: Wörter <-> Wortklassen, Textpassagen <-> Kontext (z.B. Diagnose, Familienanamnese, Labor, Medikation)
 - Nicht annotierte Korpora: z.B. Worthäufigkeiten, Wortkollokationen
 - Training probabilistischer Modelle

Terminiologiestandards: -
SNOMED CT
- LOINC

Interface-Terminologien
(anwendernah)

Ausblick

- Gering strukturierte Texte in der elektronischen Krankenakte:
 - oft einzige Informationsquelle
 - oft verlässlicher und nachhaltiger als strukturierte Inhalte
 - bisher wenige erschlossene "Goldmine"
 - Liefert wertvolle Daten für
 - Personalisierte Medizin
 - Entscheidungsunterstützung
 - Klinische Forschung ("phenotyping")
 - Vorhersage
- Breites Spektrum an semantischen / linguistischen Ressourcen erforderlich
- Diese Ressourcen müssen auf die deutsche Medizinsprache zugeschnitten werden: beträchtlicher Aufwand
- Anlehnung an internationale Standards sinnvoll. Wichtigstes Beispiel: SNOMED CT als umfassende klinische Ontologie



Beispiel: Dekomposition

Dekomposition

Ad-hoc-Komposita:

Cholangiopankreatikographie

Hausstaubmilbenphobie

Phenylbutylpiperadinderivate

Laryngotracheobronchoskopie

hypothalamikohypophyseales Erscheinungsbild

Phosphoethanolaminurie

Chorionzottenbiopsietermin

Hyperviskositätssyndrom

Dekomposition

Ad-hoc-Komposita:

Chol angio pankreat iko graph ie

Hausstaub milb en phob ie

Phenyl butyl piperadin derivate

Laryng o trache o bronch o skop ie

hypo thalamik o hypo phys eal es Erscheinung s bild

Phosph o ethanol amin urie

Chorion zotten biopsie termin

Hyper viskosität s syndrom

Dekomposition

Ad Komposit

Galle

Bauchspei-
cheldrüse

Chol angio pankreat iko graph ie

Hausstaup en phob ie

Gefäß

Bild-
gebung

Phenyl butyr piperadin derivate

Laryng o trache o bronch o skop ie

hypo thalamik o hypo phys eal es Erscheinung s bild

Phosph o ethanol amin urie

Chorion zotten biopsie termin

Hyper viskosität s syndrom

- Schulz S, Hahn U. Morpheme-based, cross-lingual indexing for medical document retrieval. Int J Med Inform. 2000 Sep;58-59:87-99
- Daumke P, Schulz S, Müller ML, Dzeyk W, Prinzen L, Pacheco EJ, Cancian PS, Nohama P, Markó K. Subword-based semantic retrieval of clinical and bibliographic documents. Methods Inf Med. 2010;49(2):141-7

Morphosemantische Analyse

Myokarditis:

#muscle #heart #inflamm;

Herzmuskelentzündung:

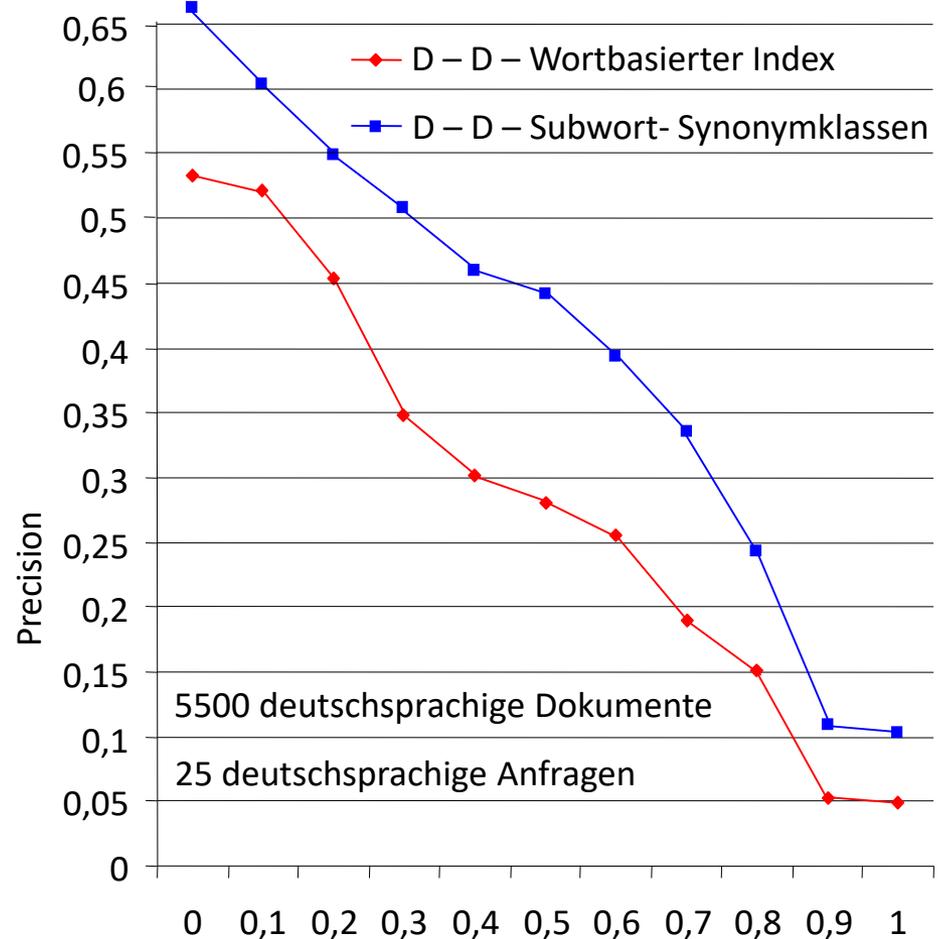
#heart #muscle #inflamm;

Inflammation of the heart muscle:

#inflamm #heart #muscle;

- Morphosemantischer Parser
- Wortgrammatik als Endlicher Automat
- Output 1: Morpheme / Subwords (ohne Semantik)
- Output 2: Subwort-Äquivalenzklassen (mit Semantik), z.B.
#heart = {"herz", "cor", "card"}_{DE}

→ Terminologie / Ontologie



- Schulz S, Hahn U. Morpheme-based, cross-lingual indexing for medical document retrieval. Int J Med Inform. 2000 Sep;58-59:87-99
- Daumke P, Schulz S, Müller ML, Dzeyk W, Prinzen L, Pacheco EJ, Cancian PS, Nohama P, Markó K. Subword-based semantic retrieval of clinical and bibliographic documents. Methods Inf Med. 2010;49(2):141-7