

Automatische Verarbeitung und Analyse deutscher klinischer Texte und Dokumente

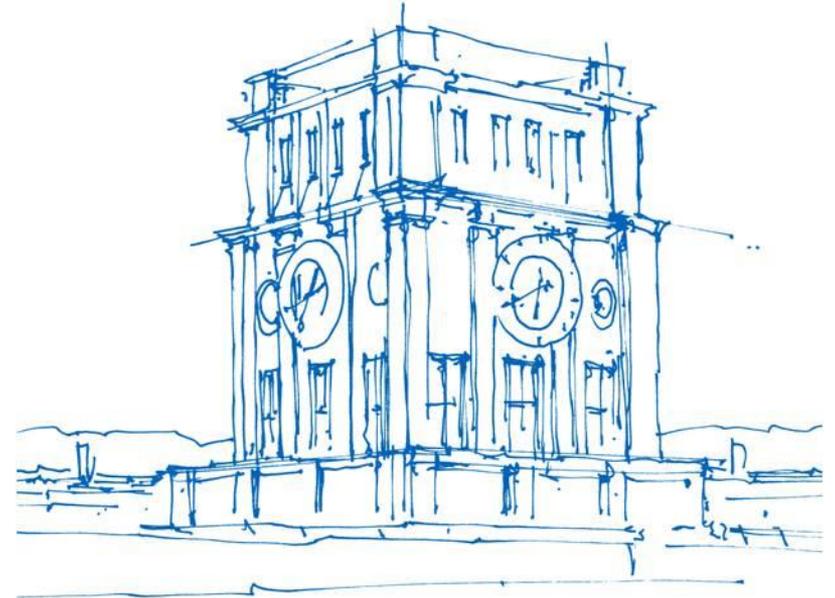
Luise Modersohn

Technische Universität München

Klinikum rechts der Isar

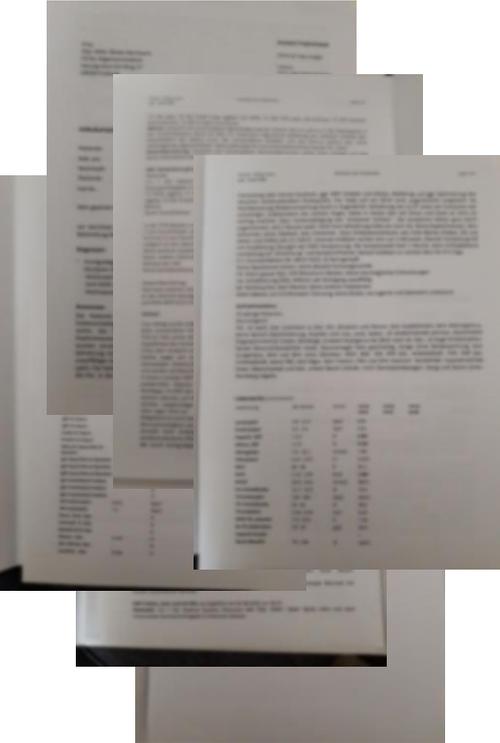
Lehrstuhl für Medizinische Informatik

Nachwuchsgruppe DE.xt



Uhrenturm der TUM

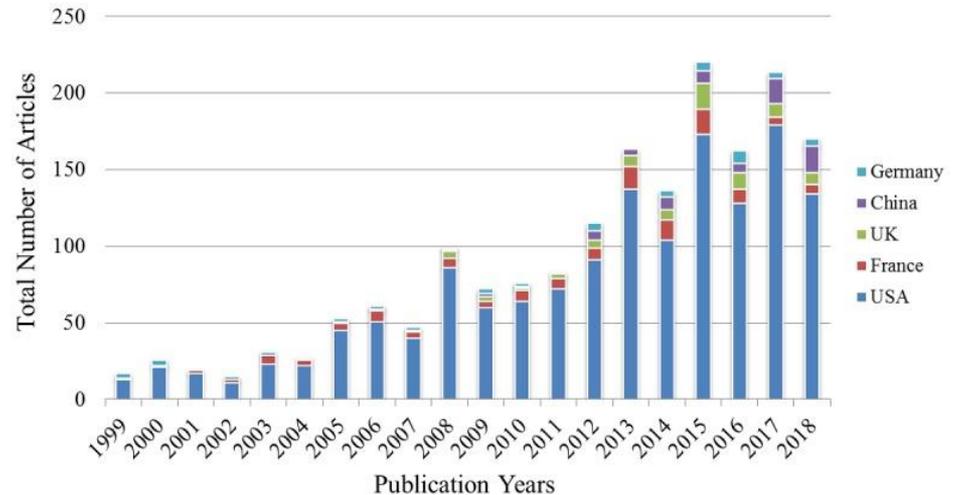
Wozu brauchen wir (klinische) Textverarbeitung



- Klinische Dokumentation in Freitextform
 - Unstrukturierte Informationen
 - Unerschlossene Datenquellen
- NLP als Werkzeug zur Überführung von Freitext zu strukturierten Daten
- Internationale Erfolge
 - Boston Children Hospital – cTAKES
 - Smith et al.; AMIA 2021
 - NLP zur Erkennung von Pneumonie bei Kindern
 - Wen et al.; npj digital medicine 2019
 - Überblick über 10 Jahre NLP an der Mayo Clinic
 - Rajkomar et al.; npj digital medicine 2018
 - Erfolg prediktiver Modelle aus EHR inklusive NLP

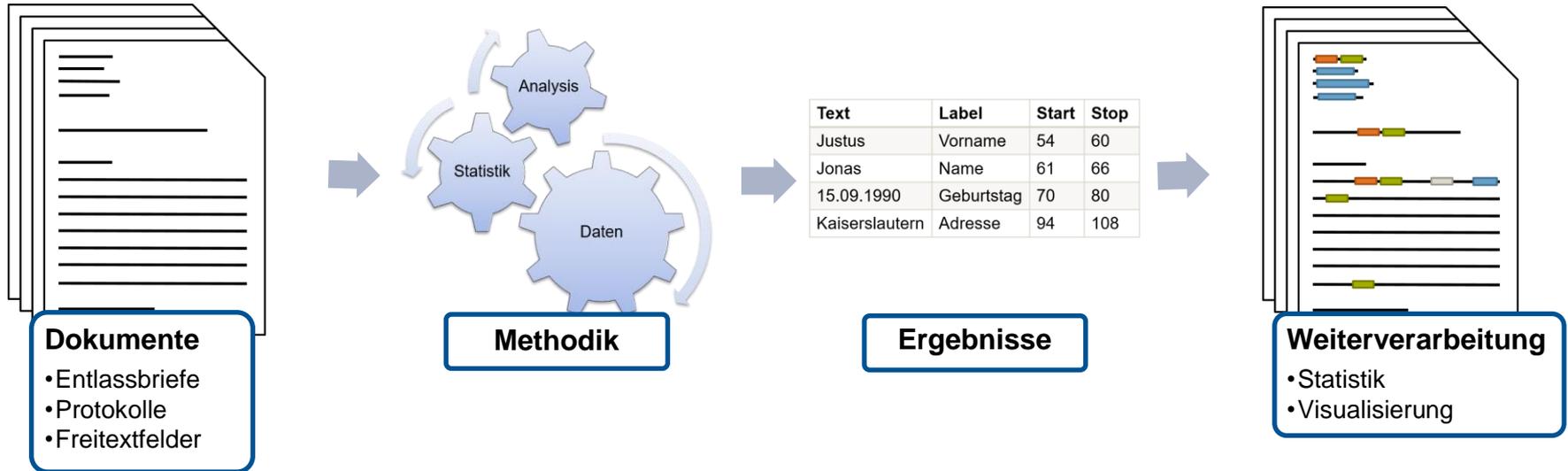
Wie ist der aktuelle Stand in Deutschland

- Zusammengesetzte Substantive
 - Methylprednisolonpulsstherapie
- Klinischer Fachjargon
 - XY wurde punktiert
- Telegramstil
 - Pupillen mw, rund, isokor
- Abkürzungen
 - Redurierter AZ, normaler EZ
- Temporaler Kontext
 - Z.n. Appendektomie



Wang et al.; Systematic Evaluation of Research Progress on Natural Language Processing in Medicine Over the Past 20 Years; JMIR 2020; Fig. 3

Vom Freitext zu strukturierten Daten



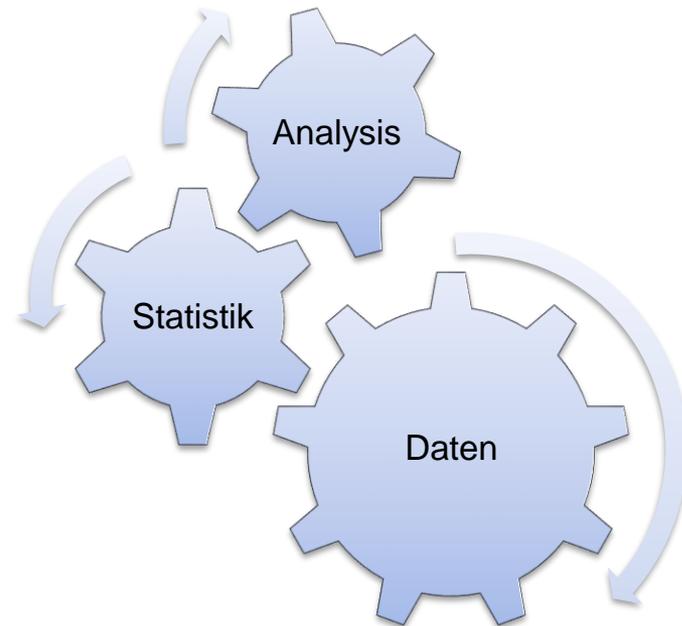
Aspekte der Methodik

Einsatzgebiete

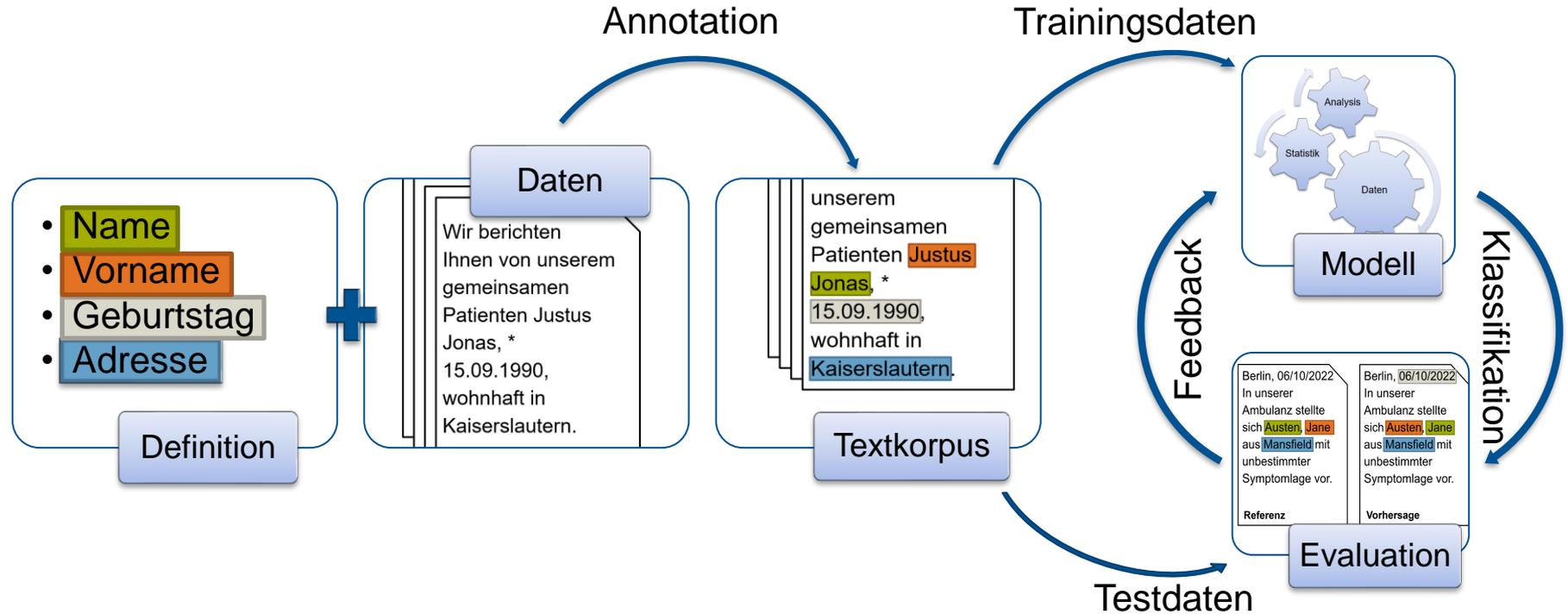
- Maschinelle Übersetzung
- Intelligente Suche
- Identifizierung wichtiger Textstellen
- Verknüpfung von Informationen

Fachgebiete

- Medizin
 - Domänenwissen
- Computerlinguistik
 - Textverarbeitung
- Informatik
 - Algorithmen/Machine Learning



Automatische Erkennung personenbezogener Elemente



Evaluation

Berlin, 06/10/2022
 In unserer
 Ambulanz stellte
 sich **Austen, Jane**
 aus **Mansfield** mit
 unbestimmter
 Symptomlage vor.

Referenz

Berlin, 06/10/2022
 In unserer
 Ambulanz stellte
 sich **Austen, Jane**
 aus **Mansfield** mit
 unbestimmter
 Symptomlage vor.

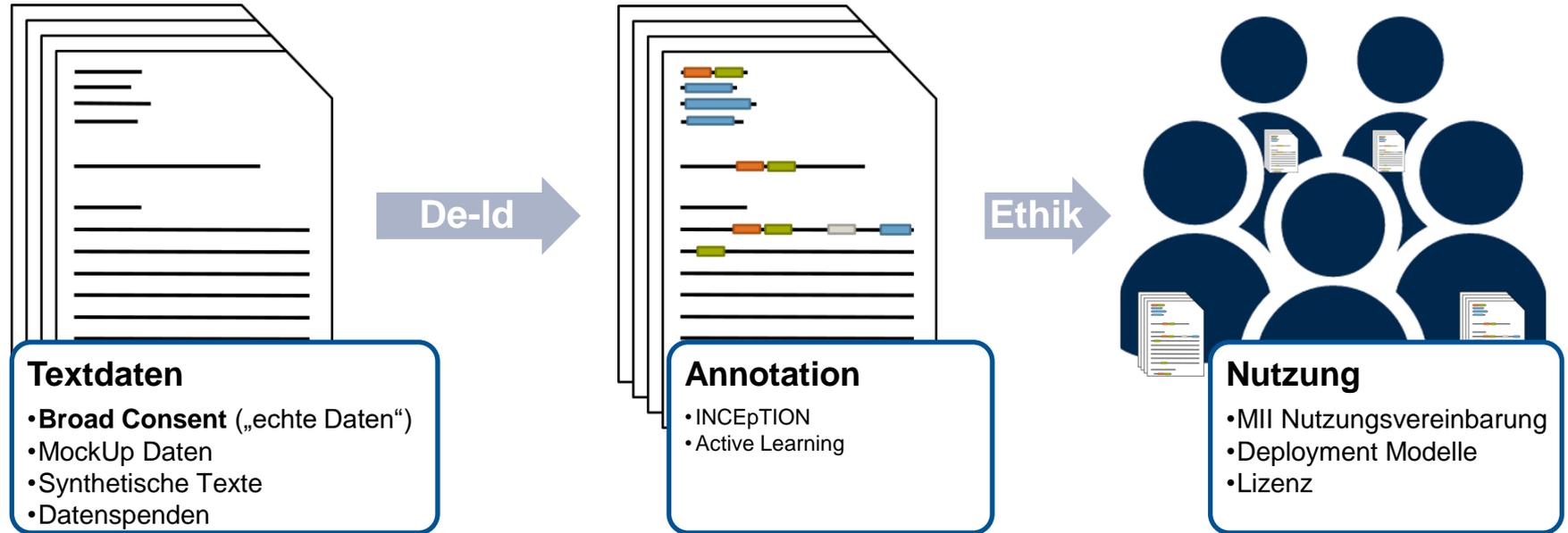
Vorhersage

Label:

- **Name**
- **Vorname**
- **Geburtstag**
- **Adresse**
- [kein Label]

Richtig ✓	12	80 %
Falsch !	3	20 %

Erstellung medizinischer Referenz-Korpora



Was wollen wir erreichen

1. Methodenentwicklung
2. Generierung von Corpora
3. Anwendung
 - Demonstration der Methodik anhand von Use Cases
 - Verbesserung der Dokumentation
 - De-Identifizierung und Pseudonymisierung
 - Verbessertes Zugriff auf Informationen
4. Nachhaltigkeit
5. Aufbau einer Community



Luise Modersohn

- Gruppenleiter
- Luise.modersohn@tum.de



Ingrid Martin

- Doktorandin
- Ingrid.martin@tum.de



Open Position

- PostDoc



Open Position

- Doktorandin

Verarbeitung deutscher klinischer Texte

- Methodenentwicklung
- Anwendung
- Vernetzung

