

Anonymisierung und Pseudonymisierung: Anforderungen der Datenverarbeiter (Entwickler)

Alfred Winter

imise.

Institut für Medizinische Informatik, Statistik und Epidemiologie

Arztbriefe mit NLP
analysieren

Warnsystem zur
Beatmung auf ITS
entwickeln

Patienten-
identifikation testen

Patienten Kontrolle
über Nutzung von
Daten einräumen

Unstrukturierte Daten aus Arztbriefen extrahieren

- Klinische Forschung benötigt viele phänotypisierte Patienten/Individuen
- Zu jedem Patienten gibt es viele phänotypische Informationen ...
... aber zu oft nur unstrukturiert auf Papier, in TIFF- oder PDF-Dokumenten.
- Natural Language Processing (NLP) kann Daten extrahieren ...
... aber die besten Verfahren funktionieren nur bei englischen Texten.

Arztbriefe mit NLP
analysieren

Warnsystem zur
Beatmung auf ITS
entwickeln

Patienten-
identifikation testen

Patienten Kontrolle
über Nutzung von
Daten einräumen

Unstrukturierte Daten aus Arztbriefen extrahieren

- *Projekt:* Entwicklung von NLP-Verfahren zur Extraktion der Medikation aus Arztbriefen.
Testdaten: je 1.000 Arztbriefe aus drei Universitäts-Klinika.
- *Problem:* Kann man Arztbriefe anonymisieren (siehe heute Nachmittag!)?
- *Lösung:* Arztbriefe von Verstorbenen?

Arztbriefe mit NLP
analysieren

Warnsystem zur
Beatmung auf ITS
entwickeln

Patienten-
identifikation testen

Patienten Kontrolle
über Nutzung von
Daten einräumen

Warnsystem zur Beatmung auf der ITS entwickeln

- Viele Patienten auf der ITS könnten von einer lungenprotektiven Beatmung profitieren ...
... wenn die Beatmung rechtzeitig eingeleitet würde.
- Ein Warnsystem ist denkbar, das die großen Datenmengen des PDMS intelligent und zeitnah auswertet.

Arztbriefe mit NLP
analysieren

Warnsystem zur
Beatmung auf ITS
entwickeln

Patienten-
identifikation testen

Patienten Kontrolle
über Nutzung von
Daten einräumen

Warnsystem zur Beatmung auf der ITS entwickeln

- *Projekt:* Entwicklung eines Warnsystems für die Einleitung lungen-protektiver Beatmung.
Testdaten: entsprechende Phänotypen aus drei Universitäts-Klinika.
- *Problem:* Wie erklärt ein Patient seine Einwilligung gegenüber einem Verbund aus drei Universitätsklinika?
- *Lösung:* Gründung eines Vereins?

Arztbriefe mit NLP
analysieren

Warnsystem zur
Beatmung auf ITS
entwickeln

Patienten-
identifikation testen

▶ Patienten Kontrolle
über Nutzung von
Daten einräumen

Patientenidentifikation für Record-Linkage in Krebsregistern testen

- In Krebsregistern mit Pseudonymen müssen Datensätze zu demselben Patienten aus unterschiedlichen Quellen zusammengeführt werden.
- Welche Pseudonymisierungsverfahren liefern die wenigsten Fehler beim Record-Linkage?

Arztbriefe mit NLP
analysieren

Warnsystem zur
Beatmung auf ITS
entwickeln

Patienten-
identifikation testen

Patienten Kontrolle
über Nutzung von
Daten einräumen

Patientenidentifikation für Record-Linkage in Krebsregistern testen

- *Publikation in der MIBE*: Test eines Psudonymisierungsverfahrens gegen einen Testdatensatz als Goldstandard.
- *Gutachterfrage*: Können die Daten des Goldstandards publiziert werden?
- *Antwort der Autoren*: Nein, Datenschutz!

Arztbriefe mit NLP
analysieren

Warnsystem zur
Beatmung auf ITS
entwickeln

Patienten-
identifikation testen

Patienten Kontrolle
über Nutzung von
Daten einräumen

Patienten die Kontrolle über die Nutzung ihrer Daten in einem Forschungskonsortium einräumen

Access Control <-> Data Usage Control

- Patient kann jederzeit sehen, wer welche seiner Daten wofür nutzt (Transparenz).
- Patient kann jederzeit die Zustimmung zur Nutzung seiner Daten zurückziehen/geben (Dynamic Consent)

Arztbriefe mit NLP
analysieren

Warnsystem zur
Beatmung auf ITS
entwickeln

Patienten-
identifikation testen

Patienten Kontrolle
über Nutzung von
Daten einräumen

Patienten die Kontrolle über die Nutzung ihrer Daten in einem Forschungskonsortium einräumen

- Erhöht die Partizipation der Patienten (Data Usage Control mit Transparenz und Dynamic Consent) die Rechtssicherheit bei der Verwendung der Patientendaten für Forschung und Entwicklung?