

TOP II Künstliche Intelligenz in der Medizin: Die Zukunft des Gesundheitswesens aus ärztlicher Perspektive gestalten

Titel: Schutz vor Reidentifizierung bei Gesundheitsdaten aus der elektronischen Patientenakte über Kalendermuster - gerade durch die Anwendung von Künstlicher Intelligenz

Beschlussantrag

Von: Dr. Stefan Streit als Abgeordneter der Ärztekammer Nordrhein
 Wieland Dietrich als Abgeordneter der Ärztekammer Nordrhein
 Dirk Paulukat als Abgeordneter der Landesärztekammer Hessen
 Dr. Christel Kreuzer als Abgeordnete der Ärztekammer Nordrhein
 Mira Faßbach als Abgeordnete der Ärztekammer Nordrhein
 Dr. Stefan Schröter als Abgeordneter der Ärztekammer Nordrhein

Der 129. Deutsche Ärztetag 2025 fordert vom Bundesgesundheitsministerium (BMG), der gematik, dem Bundesamt für Sicherheit in der Informationstechnik (BSI) und der Bundesdatenschutzbeauftragten (BfDI) vor dem Beginn der Dateneinspeisung eine detaillierte Darstellung, ob und wie die Daten aus der elektronischen Patientenakte (ePA) im Forschungsdatenzentrum (FDZ) gegen eine Reidentifizierung geschützt sind. Neben den Fragen zu Schwachstellen in der Prozessplanung und der Programmierung sind solche zur kryptografischen Sicherheit gegenüber einer Reidentifikation aufgekommen. Soll Künstliche Intelligenz (KI) zur Auswertung dieser Daten zur Anwendung kommen, bedarf es eines weitergehenden Schutzkonzepts.

Begründung:

Durch die breite Verfügbarkeit von großer Rechnerleistung gelang schon in der Vergangenheit über den Abgleich zwischen öffentlich verfügbaren, personenbezogenen Daten und anonymisierten Gesundheitsdaten deren Reidentifizierung. Durch den Einsatz von KI steigt das Risiko, dass durch einen Musterabgleich die Anonymisierung gebrochen wird. Wirksame Anonymisierung ist die Basis der Legitimation zur Verwendung von Real-World-Gesundheitsdaten in der Datenökonomie.

Ein vorab befragter Experte geht davon aus, dass für eine Reidentifizierung nur die drei Attribute Geschlecht, Alter und Postleitzahl sowie die Kenntnis der Behandlungstage von fünf Arztkontakten ausreichend sind.

Der Experte bezieht sich auf eine Studie, in der die Gesundheitsakte eines US-Gouverneurs

ANGENOMMEN

Angenommen: ☒ Abgelehnt: ☐ Vorstandsüberweisung: ☐ Entfallen: ☐ Zurückgezogen: ☐ Nichtbefassung: ☐

Stimmen Ja: 162

Stimmen Nein: 24

Enthaltungen: 17

auf Basis eines öffentlichen Wählerverzeichnisses reidentifiziert werden konnte, auf Basis von Geburtsdatum, Geschlecht und Postleitzahl.

Zusätzlich tragen Kalendermuster aus Metadaten zu einer Reidentifizierung bei. Die Wahrscheinlichkeit, dass zwei Patienten in einem Jahr an genau denselben vier Tagen zu einem Arzt gehen, liegt weit unter eins zu 82 Millionen, also der Zahl der Menschen, die in Deutschland leben. Hier entstehen Kalendermuster. Alle Menschen, die in Deutschland leben, gehen durchschnittlich zehnmal im Jahr zum Arzt, davon etwa fünfmal zum Hausarzt. Die Zahl der Termindaten steigt durch Wiederholungsrezepte ohne Arztkontakt deutlich über zehn pro Jahr.

Wie bekannt ist, verringert man das Risiko der Reidentifizierung durch Aggregation, Swapping, Verrauschen sowie der Erzeugung synthetischer Gesundheitsdaten. Diese Methoden reduzieren oder verändern aber die Nutzdaten. Der dadurch entstehende Schutz vor Reidentifizierung ist allerdings geringer, als man üblicherweise annimmt. Selbst ausprobieren kann man das mit frei verfügbaren Internet-Tools, wie z. B. VisualAnon. Gleichzeitig werden Gesundheitsdaten durch die Schutzmaßnahmen den Real-World-Daten immer unähnlicher. Deshalb müssen diese Schutzverfahren gegen die Datenverwendbarkeit in der Forschung abgewogen werden. Während konventionelle Wissenschaft mit dieser Datenveränderung vergleichsweise gut zurechtkommt, sind die Einflüsse auf KI-Erkenntnisse unkalkulierbar.

Reidentifizierung anonymisierter Gesundheitsdatensätze ist keine Theorie, sondern wurde in Amerika und Australien bereits erfolgreich demonstriert.

Während die Sicherheit gegenüber einer Reidentifizierung von Gesundheitsdaten durch Angriffe mit einem konventionellen Datenbankabgleich mathematisch kalkuliert werden kann, ist die Berechnung der Sicherheit vor Angriffen bei KI-Anwendungen nicht möglich. Das heißt, zum bereits bestehenden kryptografisch-kalkulierbaren Risiko kommt zukünftig ein unkalkulierbares KI-Risiko dazu.

Wird eine KI zuerst an Big-Data aus dem Forschungsdatenzentrum trainiert, sind die Kalendermuster der ePA-Patientendatensätze in der KI enthalten. Wird die gleiche KI anschließend mit Informationen nach § 370a SGB V und mit Rezeptdaten angelernet, dann ist davon auszugehen, dass über die Kalendermustererkennung eine Verknüpfung mit dem Namen des Patienten stattfindet. Dazu bedarf es keines expliziten "Auftrags", denn KI arbeitet immer mit Mustererkennung.

Egal zu welchem Zweck diese KI anschließend eingesetzt wird, sie greift dann auf die enthaltene Information über die Attribute aus den Gesundheitsinformationen des konkreten Patienten zu und macht Vorhersagen, die diese Informationen einbeziehen.

Problematisch ist die unklare Verantwortlichkeit durch die Vielzahl der Datenspeicherplätze. Daten, die eine Reidentifizierung ermöglichen, liegen in Arztpraxen, in Apotheken, in der

ePA, bei den gesetzlichen Krankenkassen und privaten Krankenversicherungen, bei den Institutionen, die die ärztliche Terminvergabe (auch nach § 370a SGB V) organisieren, im Forschungsdatenzentrum (mit der Zuordnungsdatei zwischen Klaridentität und dem einjährigen Forschungsdatenpseudonym) und überall dort, wo der eigene Name mit Geburtsdatum, der Adresse und anderen Nutzdaten (Gewicht, Körpergröße, Vorerkrankungen usw.) in Abfragen und (Gesundheits-)Apps vom Patienten schon einmal eingegeben wurde.

Diese Zusammenhänge werfen Fragen bezüglich der ärztlichen Schweigepflicht nach § 203 StGB auf. Wenn allgemein bekannt ist, dass bereits ohne KI eine Reidentifizierung von ePA-Datensätzen erfolgen kann, dann steht die Ärzteschaft vor einer Datenweitergabe für eine KI-Anwendung in der Verantwortung.

Unabhängig von der strafrechtlichen Bewertung zerstört eine drohende, benachteiligende Diskriminierung von Patientinnen und Patienten auf Basis öffentlich gewordener ärztlicher Behandlungsdaten die bisherige Grundlage der ärztlichen Berufsausübung.

ANGENOMMEN