TOP 1 **Aktuelle Themen der Berufs- und Gesundheitspolitik**

Titel: Schutz vor Reidentifizierung bei anonymisierten/ pseudonymisierten Gesundheitsdaten aus der elektronischen Patientenakte (ePA) im Gesundheitsdatenraum (EHDS) über (Kalender-)Mustermatching und die Anwendung von Künstlicher Intelligenz (KI)

Beschlussantrag

Von: Fraktion FuturMed29

BESCHLUSSVORSCHLAG:

Die Kammerversammlung der Ärztekammer Nordrhein fordert vom Bundesgesundheitsministerium (BMG), der gematik, dem Bundesamt für Sicherheit in der Informationstechnik (BSI) und der Bundesdatenschutzbeauftragten (BfDI) vor dem Beginn der Datenbereitstellung für die Datenökonomie eine detaillierte Darstellung, ob und wie die anonymisierten und pseudonymisierten Daten aus der elektronischen Patientenakte (ePA) im Gesundheitsdatenraum (EHDS) 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. Nach der Digitalisierungsstrategie 2026 soll Künstliche Intelligenz (KI) zur Auswertung dieser Daten zur Anwendung kommen, deshalb 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. Nur eine wirksame Anonymisierung legitimiert der 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-Governeurs 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 84 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 Reidentifizierug 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 diese Schutzmaßnahmen den Real-World- Daten immer unähnlicher. Deshalb müssen Schutzverfahren immer gegen die Datenverwendbarkeit abgewogen werden. Während konventionelle Wissenschaft mit diesen Datenveränderungen 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 KIRisiko dazu. Wird eine KI zuerst an anonymisierten oder pseudonymisierten Big-Data aus der elektronischen Patientenakte trainiert, sind die Kalendermuster der ePA-Patientendatensätze in der KI enthalten. Wird die gleiche KI anschließend an anderen Datenbanken angelernt, dann ist davon auszugehen, dass über die Kalendermusterkennung eine Verknüpfung mit dem Namen des Patienten stattfindet.

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

einer Verarbeitung mit und ohne KI-Anwendung, im Gesundheitsdatenraum, 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.