

ePA-Daten für die Forschung: Wie können sie vertrauensvoll geschützt und genutzt werden?

gematik entwickelt Demonstratoren in Kooperation mit Universität Mannheim und IBM Deutschland

Berlin, 05.05.2022 – Eine valide Datenbasis ist Voraussetzung für eine gute Gesundheitsforschung. Doch wie können Daten geschützt und vertraulich verarbeitet werden? Die gematik hat dazu mit IBM Deutschland und der Universität Mannheim Demonstratoren entwickelt. Das Thema ist aktuell wie nie: So forderte der Corona-Expertenrat, dass Deutschland eine umfassende Auswertung und Veröffentlichung anonymisierter Gesundheitsdaten benötige. Eine wichtige Rolle soll dafür die elektronische Patientenakte (ePA) spielen, denn Versicherte können Daten für die Gesundheitsforschung freigeben. Ziel ist, die Datenbasis für die Forschung und damit Heilungschancen zu verbessern. Daten müssen dabei vertrauensvoll analysiert werden, ohne dass Rückschlüsse auf Versicherte möglich sind. Eine Pseudonymisierung ihrer Daten ist hierbei oft alleine kein wirksamer Schutz. Deshalb werden alternative datenschutzfreundliche Technologien, sogenannte Privacy Enhancing Technologies (PET), die diesen Schutz gewährleisten, erforscht und auch schon an vielen Stellen eingesetzt.

Die gematik hat in Kooperation mit IBM Deutschland und Prof. Dr. Frederik Armknecht, Inhaber des Lehrstuhls für Praktische Informatik IV an der Universität Mannheim den Prototyp PrETTI (Privacy Enhancing Technologies in der Telematikinfrastruktur) gebaut. PrETTI demonstriert, wie Versicherte ihre ePA-Daten mittels moderner datenschutzfreundlicher Techniken bereitstellen können. Professor Armknecht entwickelte hierbei die Backend-Systeme zum datenschutzfreundlichen Machine Learning Verfahren, IBM Consulting passte die ePA-Oberfläche für Versicherte (ePA-Frontend des Versicherten) an.

Das Projekt PrETTI zeigt, dass es bereits heute PET gibt, die für einen praktikablen Einsatz in der Telematikinfrastruktur genutzt werden können. Die Einsatzmöglichkeiten in der TI sollen nun evaluiert und zusammen mit weiteren Partnern ausgebaut werden.

Videos:

<https://youtu.be/c0vHWkd6qY>

<https://youtu.be/dHwe03MdL1I>

Hintergrund:

PrETTI betrachtet zwei Szenarien:

1) Maschinelles Lernen mit Multi-Party-Computation: In diesem Szenario gibt der Versicherte an seinem ePA-FdV Daten für ein Forschungsvorhaben frei (z.B. für die Forschung zu Schlaganfällen). Die freigegebenen Daten werden im ePA-FdV so bearbeitet, dass daraus pro Datum mehrere Teile entstehen, die einzeln für sich mathematisch nachweisbar keinen Rückschluss auf das ursprüngliche Datum mehr ermöglichen. Diese Teile werden als shares bezeichnet. Diese shares werden dann an mehrere voneinander unabhängige Parteien (Dienste) verteilt, die damit mittels Multi-Party-Computation ein Machine Learning Modell zu einer Forschungsfrage lernen. Die einzelnen Parteien nutzen dabei nur die shares und haben niemals Kenntnis über die ursprünglichen Daten des Versicherten. Auch das gelernte Modell gibt keinen Rückschluss über die Versicherten.

2) Anwendung des Modells mittels Homomorphic Encryption: Wurde ein Modell zu einer Forschungsfrage gelernt, kann der Versicherte dieses über sein ePA-FdV nutzen. Wurde z.B. ein Modell zum Schlaganfallrisiko gelernt, können Versicherte sich am ePA-FdV ihr Schlaganfallrisiko einschätzen lassen. Hierzu müssen die Versicherten dem Modell natürlich entsprechende Gesundheitswerte übergeben. Damit aber auch bei der Modellanwendung keine Rückschlüsse auf die Versicherten gezogen werden können, hat IBM Consulting im ePA-FdV Homomorphic Encryption implementiert. Dies ermöglicht die Daten der Versicherten im ePA-FdV so zu verschlüsseln, dass die Modellanwendung ausschließlich auf verschlüsselten Daten erfolgt und keine Klartextdaten benötigt werden. Das Auswertungsergebnis ist ebenfalls verschlüsselt. Es wird erst im ePA-FdV entschlüsselt und dort dem Versicherten angezeigt.

Kontakt: Pressestelle gematik GmbH – Tel. +49 (0) 30 40041-441 – presse@gematik.de