Lernende Systeme
DIE PLATTFORM FÜR KÜNSTLICHE INTELLIGENZ

# Sichere KI-Systeme für die Medizin

Whitepaper von Jörn Müller-Quade et al.
Arbeitsgruppe IT-Sicherheit, Privacy, Recht und Ethik

## Kurzfassung

Der Einsatz von Künstlicher Intelligenz (KI) verspricht in der Medizin große Verbesserungen. Lernende Systeme können künftig bei der Prävention, frühzeitigen Diagnose sowie der patientengerechten Therapie zu besseren Behandlungsergebnissen führen und somit unsere Gesundheitsfürsorge verbessern. Der Einsatz von KI-Systemen kann zudem Ärztinnen und Ärzte sowie medizinisches Pflegepersonal bei der Patientenversorgung unterstützen und das medizinische Personal entlasten. Konkrete Beispiele liefert das fiktive Anwendungsszenario „Mit KI gegen Krebs", das von der Arbeitsgruppe Gesundheit, Medizintechnik, Pflege der Plattform Lernende Systeme entwickelt wurde.

Gleichzeitig stellt der Einsatz von intelligenten und selbstlernenden Systemen im Gesundheitswesen hohe Anforderungen an das Datenmanagement und die IT-Sicherheit der Systeme. Mögliche Risiken sind unter anderem fehlerhaft oder bewusst verfälschte Trainingsdaten, Angriffe auf die KI-Software oder die fehlerhafte Integration in die klinische Praxis. Entlang des genannten Anwendungsszenarios haben Mitglieder der Arbeitsgruppe IT-Sicherheit, Privacy, Recht und Ethik sowie der Arbeitsgruppe Gesundheit, Medizintechnik, Pflege der Plattform Lernende Systeme Anforderungen an die IT-Sicherheit identifiziert, die für den Einsatz von KI-Assistenzsystemen in der Medizin notwendig sind.

### Anforderungen an den Einsatz von KI in der Medizin

- Originale, unverfälschte Trainingsdaten sicherstellen
- KI-Software vor Angriffen schützen
- Trainingsdaten unter Wahrung der Privatsphäre poolen
- Sichere KI-Datenbanken
- Patientendaten sicher bereitstellen
- KI-Systeme sicher in den klinischen Prozess integrieren

Die Experten fokussieren dabei das Datenmanagement und Sicherheitsaspekte – und damit primär technische Fragen. Diese technische Analyse ist ein grundlegender Schritt, um regulatorische Fragen diskutieren und beantworten zu können. KI und Maschinelles Lernen (ML) werfen auch gesellschaftsrelevante Fragestellungen auf, die sich nicht rein technisch beantworten lassen. Das gilt insbesondere für das Gesundheitswesen, wo die benötigten Daten besonders sensibel und mögliche Risiken schwerwiegend sind.

Das Whitepaper identifiziert dazu technische und organisatorische Bedingungen, die für eine mittelfristige Realisierung des von der Plattform Lernende Systeme entwickelten fiktiven Anwendungsszenarios „Mit KI gegen Krebs" notwendig sind. Mit einer technisch-organisatorisch fundierten Analyse soll die Grundlage für eine Folgediskussion gelegt werden. Denn die weitere Gestaltung des digitalisierten Gesundheitswesens ist auch relevant für das zugrunde liegende Anwendungsszenario. Folgende Aspekte wurden daher hier offengelassen und müssen Gegenstand künftiger Diskussionen sein:

- Zugriffsberechtigung für Dritte auf die Einträge der elektronischen Patientenakte (ePA)
- Freiwillige und geschützte Datenfreigabe
- Prüfbedarf der Rechtslage

Die angesprochenen Aspekte sind aktuell noch rechtlich ungeklärt. Ausgehend von der Analyse des Anwendungsszenarios formulieren die Experten rechtliche Gestaltungserfordernisse und mögliche Gestaltungsoptionen. Dabei liegt der Fokus auf der Frage der Qualitätsabsicherung der für das Training von KI-Systemen verwendeten Daten, der Nachvollziehbarkeit und Erklärbarkeit von KI-Systemen sowie deren Sicherheit im Sinne von Safety und IT-Security.

## Rechtlich-regulatorische Erfordernisse und mögliche Gestaltungsoptionen

- **Gemeinsame Leitlinien und Prüfvorschriften für die Zulassung und Zertifizierung entwickeln:** Dynamische Softwarearchitekturen führen dazu, dass die Funktion und Wirkungsweise eines Medizinproduktes vor Inverkehrbringung weniger mess-, beleg- und im Bedarfsfall zertifizierbar ist. Dies trifft auch auf Lernende Systeme zu. Daraus abzuleiten, dass ein Produkt bei jedem geringfügigen Software-Update als neues Produkt zu betrachten wäre, erscheint nicht zielführend. Gleichwohl sollte der Zulassungsprozess weiterentwickelt werden. Neben dem Produkt selbst ist es auch notwendig, dessen Betrieb und Zertifizierungsanforderungen für Updates zu betrachten.
- **Gemeinsame Leitlinien und Prüfvorschriften für die Zulassung- und Zertifizierung der Betreiber der KI-Datenbanken entwickeln:** Der Gesetzgeber sollte gemeinsam mit den relevanten Stakeholdern ebenfalls Leitlinien sowie Prüfvorschriften und Anforderungen an einen Zulassungs- und Zertifizierungsprozess für die zertifizierten Betreiber der KI-Datenbanken entwickeln.
- **Hersteller gesetzlich zur Mängelbehebung verpflichten:** Es entstehen neue, ggf. strengere Sicherheitsanforderungen an die Anwendungen, die im Rahmen der Zulassung erfüllt werden müssen. Für den Betrieb eines KI-Systems ist dies in der europäischen Gesetzgebung geregelt und fest definiert. Bestimmte Produkteigenschaften können vor Markteinführung überprüft und bewertet werden. Darüber hinaus sollten nachgelagert auch Funktionsstörungen beobachtet und von den Herstellern im Sinne der Mängelbehebung behoben werden – unabhängig davon, ob sie nur auf KI-Funktionalitäten oder sonstige Systemanpassungen zurückzuführen sind.

- **Unabhängige autorisierte Betreiber des KI-Assistenzsystems einsetzen:** Diese staatlich beauftragten neutralen Einrichtungen sollte damit beauftragt werden, die Analyseverfahren und Datensätze zu verwalten und zu pflegen. Diese Einrichtung darf nicht befugt sein, Daten zu verändern oder einzuspeisen, da ein eigenes ökonomisches Interesse vorliegen könnte.
- **Ein unabhängiges Prüf-Komitee einsetzen:** Ein interdisziplinäres Expertengremium sollte in regelmäßigen Abständen die Funktionsweise der zertifizierten und eingesetzten KI-Systeme überprüfen. Es wäre sinnvoll, dieses Komitee beim Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM) einzurichten. Ferner sollten bei den Herstellern Rückrufprozesse etabliert werden, um im Falle des Nicht-Funktionierens eines Systems handlungsfähig zu sein.
- **Krankenkassen sollten Sperrlisten führen:** Als ausgebende Stellen der elektronische Gesundheitskarten (eGKs) und der Heilberufsausweise (HBAs) sollten die Krankenkassen Sperrlisten führen, um einen unautorisierten Zugriff auf Daten zu verhindern. Diese Listen müssen fortlaufend aktualisiert werden, sodass im Fall des Verlusts die jeweilige Berechtigungskarte wertlos ist. Der Sperr-Notruf 116 116 könnte um die eGK und die HBA erweitert werden. Zu erwägen ist auch eine entsprechende Selbstverpflichtung der Krankenkassen, sich an dem System zu beteiligen.
- **Rückfall-Lösung einführen:** Eine Rückfall-Lösung könnte die Sperrung der eGK ergänzen. Bei ihr handelt es sich um einen Modus, in dem der Funktionsumfang eingeschränkt ist, aber die wichtigsten Funktionen eines Systems aufrechterhalten bleiben können.
- **Mindestanforderungen an sichere Dateninfrastrukturen und Rechenzentren formulieren:** Die IT-Infrastrukturen, die für eine Umsetzung des Anwendungsszenarios gebraucht werden, unterliegen aktuell bereits dem Anwendungsbereich geltender Rechtssetzung. Dies schließt allerdings nicht aus, dass der Gesetzgeber die einschlägigen Rechtsverordnungen anpassen kann, indem er Mindestanforderungen definiert. Diese sollten festlegen, dass die Daten nur innerhalb der Europäischen Union gespeichert und erarbeitet werden dürfen. Dabei ist in einem ersten Schritt wichtig, dass die verwendeten KI-Systeme und die damit verbundenen Infrastrukturen auch von der BSI-Verordnung für kritische Infrastrukturen (KRITIS-VO) systematisch erfasst werden. In einem zweiten Schritt sind auch entsprechende Sicherheitsanforderungen für den Aufbau der notwendigen KI-Systeme festzulegen.
- **Eine forschungskompatible elektronische Patientenakte (ePA) einführen:** Damit Patientinnen und Patienten ihre Datensätze nach der Behandlung der (universitären) Forschung zur Verfügung stellen und KI-Methoden weiterentwickelt werden können, bedarf es einer forschungskompatiblen ePA. Das bedeutet, dass die relevanten Daten in einer hohen Qualität, vollständig und in einer weiterverwendbaren Form vorliegen sollten.
- **Elektronische Patientenakte (ePA) zu einer erweiterten elektronischen Patientenakte (eePA) erweitern:** Vor allem bei der Vorsorge werden weitere Patientendaten benötigt, um Patientinnen und Patienten statistisch verlässlich zu möglichen Risikogruppen zuordnen zu können.
- **IT-Sicherheitsprobleme weiter erforschen:** Nicht alle beschriebenen IT-Sicherheitsprobleme, die beim Einsatz von KI-Systemen im Gesundheitswesen auftreten könnten, können mit den aktuell zur Verfügung stehenden technischen Lösungen beantwortet werden. Deshalb ist die Wissenschaft aufgerufen, diese Probleme zu erforschen und möglichst zuverlässige Lösungen zu entwickeln. Dafür sollten passende Programme ins Leben gerufen und die entsprechende Forschungsförderung bereitgestellt werden.

Verknüpft werden diese möglichen Gestaltungsoptionen mit gesellschaftsrelevanten Fragestellungen, etwa zu Nutzen und potentiellen Risiken des Einsatzes von KI-Systemen im Gesundheitswesen und zur Verwendung von anonymisierten oder pseudonymisierten Daten. Diese müssen in einem gesellschaftlichen Diskurs erörtert und beantwortet werden.

## Gesellschaftsrelevante Fragestellungen zum Einsatz von KI in der Medizin

- **Dateninfrastruktur betreiben, warten und pflegen:** Patientinnen und Patienten können mithilfe der eGK souverän über ihre Daten entscheiden. Die daran angeschlossene ePA bildet eine Schnittstelle zwischen Patientinnen und Patienten, behandelnden Ärztinnen und Ärzten und den KI-Systemen. Abschließend geklärt ist nicht, wo und in welcher Form Daten (zwischen-) gespeichert, übertragen und erweitert werden. Dies betrifft sowohl die elektronischen Patientendaten als auch die Meta-Daten der ePA, die die KI-Software bewertet hat. Verteilte Cloud-Infrastrukturen könnten einen Lösungsansatz darstellen, da diese im Rahmen bestehender KRITIS-Regulierungen bereits größtenteils abgedeckt sind. Offen ist, wer die dafür notwendige Infrastruktur bereitstellt, sie wartet und pflegt.
- **KI-Assistenzsystem bereitstellen und betreuen:** Zu klären sind außerdem Fragen der operativen Umsetzung, also etwa, welche Institutionen die KI-Systeme finanzieren, pflegen, kontinuierlich trainieren und auf Anfrage einer Ärztin oder eines Arztes die neueste KI-Software zur Verfügung stellen können. Diese Institutionen müssen unabhängig sein und dürfen über keine eigenen Einspeise- oder Veränderungsmöglichkeiten verfügen.
- **Nutzen und Risiko abwägen:** Wie viele andere medizinische Methoden der Diagnostik und Therapie bergen auch KI-Assistenzsysteme gewisse Risiken. In der Diagnostik können falsch-positive und falsch-negative Ergebnisse zu Fehlbehandlungen und schweren physischen, psychischen und finanziellen Belastungen führen. Derartige Risiken werden sich nicht vollständig ausschließen lassen. Mit KI könnten neue Nutzen-Risiko-Abwägungen notwendig werden. So könnten mithilfe von Big-Data-Analysen insgesamt mehr Krankheiten früher entdeckt werden. Dies könnte aber auch mit dem Risiko von mehr falsch-positiven Befunden einhergehen. Deshalb sollte in einem gesellschaftlichen Diskurs die Frage geklärt werden, unter welchen Umständen und bis zu welcher Höhe sind wir bereit, als Gesellschaft „Fehlerquoten" zu akzeptieren, wenn auf der anderen Seite hoher medizinischer Nutzen geschaffen werden kann?
- **Verwenden von Daten (Zweckgebundenheit):** Wie Patientinnen und Patienten den Zugriff auf ihre Daten und deren weitere anonymisierte beziehungsweise pseudonymisierte Verwendung (z. B. für Forschungsprojekte) autorisieren sollten und können, bedarf weiterer Ausgestaltung und Konkretisierung. Es muss daher geklärt werden, welche Daten sie weitergeben können und wie eng die Zweckgebundenheit einer freiwilligen und geschützten Datenweitergabe bei explorativer Forschung ausgelegt werden sollte.
- **Verantwortung und Haftung:** Grundsätzlich sollte der Mensch Letztentscheider sowohl für den Behandlungsablauf als auch für den Umgang mit seinen Daten bleiben. Aber auch dann könnten falsch verarbeitete Informationen möglicherweise schwerwiegende Behandlungsfehler verursachen, etwa bei einer Operation. Zu diskutieren ist, wer die Verantwortung für Fehler trägt und ob der Einsatz von KI-Systemen im Sinne einer Haftung versicherbar sein sollte. Daraus folgt die Frage, wie die Verantwortung und Haftung zwischen dem Anbieter und Betreiber des KI-Systems und dem medizinischen Personal aufgeteilt werden sollte.

- **Transparenz der Ergebnisse, Nachvollziehbarkeit versus Erklärbarkeit:** Je komplexer ein KI-Verfahren, desto intransparenter werden die Berechnungsschritte, mithilfe derer es zum Ergebnis kommt. Dies birgt die Gefahr, dass Anwenderinnen und Anwender korrekte Ergebnisse falsch interpretieren. Ebenso denkbar ist, dass sie verzerrte oder manipulierte Ergebnisse unbemerkt verwenden. Für eine korrekte Behandlung ist es deshalb wichtig zu wissen, warum ein Ergebnis ausgegeben wurde (Kausalität). So wünschenswert eine maximale Nachvollziehbarkeit einerseits erscheint, könnte sie andererseits zu einer Informationsüberflutung führen. Dieses Spannungsfeld gilt es im gesellschaftlichen Diskurs auszutarieren. Daraus folgt die Frage, wieviel Auskunftsrecht über die Berechnung eines KI-Systems Ärztinnen, Ärzte sowie Patientinnen und Patienten haben müssen. Außerdem sollte geklärt werden, welche Regeln der Gesetzgeber für die Nachvollziehbarkeit und Erklärbarkeit von KI-basierten Medizinprodukten schaffen sollte.

**Impressum**
Herausgeber: Lernende Systeme – Die Plattform für Künstliche Intelligenz | Geschäftsstelle | c/o acatech | Karolinenplatz 4 | D-80333 München | kontakt@plattform-lernende-systeme.de | www.plattform-lernende-systeme.de |
Folgen Sie uns auf Twitter: @LernendeSysteme | Stand: April 2020 | Bildnachweis: Tom Werner/gettyimages/Titel

Diese Kurzfassung entstand auf Grundlage des Whitepapers Sichere KI-Systeme für die Medizin Datenmanagement und IT-Sicherheit in der Krebsbehandlung der Zukunft, München, 2020. Es wurde erstellt von den Arbeitsgruppen IT-Sicherheit, Privacy, Recht und Ethik unter Mitwirkung der Arbeitsgruppe Gesundheit, Medizintechnik, Pflege. Die Originalfassung der Publikation ist online verfügbar unter: https://www.plattform-lernende-systeme.de/publikationen.html

GEFÖRDERT VOM

Bundesministerium für Bildung und Forschung

acatech
DEUTSCHE AKADEMIE DER TECHNIKWISSENSCHAFTEN