



Beschluss des Stadtrats

vom 9. September 2026

GR Nr. 2026/282

Nr. 3171/2026

Schriftliche Anfrage von Vera Çelik und Yves Henz betreffend Einsatz von Künstlicher Intelligenz (KI) im Gesundheits- und Umweltsdepartement (GUD) und im Stadtspital Zürich, Analyse der eingesetzten und geplanten KI-Systeme inklusive Anbieter und Datenquellen, Prüfung der Datenverarbeitung innerhalb und ausserhalb der städtischen IT-Infrastruktur, Risiken durch ausländische Cloud-Anbieter und Zugriff durch US-Behörden gemäss CLOUD Act, Sicherstellung von Datenschutz, algorithmischer Fairness und Transparenz für BIPOC- und FINTA-Personen sowie Haltung zur Vermeidung algorithmischer Diskriminierung und zur Pflicht zur Information von Menschen in medizinischer Versorgung

Am 10. Juni 2026 reichten die Mitglieder des Gemeinderats Vera Çelik (SP) und Yves Henz (Grüne) folgende Schriftliche Anfrage, GR Nr. 2026/282, ein:

In seiner Antwort vom 25. September 2024 (GR Nr. 2024/343) hat der Stadtrat die Chancen und Risiken von Künstlicher Intelligenz (KI) im Gesundheits- und Umweltsdepartement (GUD), insbesondere am Stadtspital dargelegt. Die damaligen Antworten vom Stadtrat liessen jedoch in einigen Hinsichten einige Kernfragen offen.

Der Stadtrat erläuterte, dass die führenden KI-Systeme, meist als Cloud-Dienste ausländischer Anbieter laufen, oft isolierte Inselösungen sind und massive Investitionskosten verursachen. Auch die rechtliche Haftung und der Datenschutz sind ungeklärt. Zudem räumte er ein, dass KI Entscheidungen oft eine undurchsichtige «Blackbox» bleiben und somit das Risiko für rassistische und sexistische Verzerrungen (Bias) im Patient*innengut bestehen bleibt.

Wenn klinische KI -Systeme auf ausländischen Cloud-Infrastrukturen operieren, droht ein schleichender Verlust der digitalen Souveränität der Stadt Zürich. Gleichzeitig besteht das akute Risiko, dass KI-Modelle, die auf historisch verzerrten Daten basieren, bestehende Diskriminierungen im Gesundheitssystem replizieren und verstärken. Betroffen sind hiervon insbesondere BIPOC (Black, Indigenous, People of Color), FINTA (Frauen, intergeschlechtliche, nicht-binäre, trans- und agender Personen) sowie armutsbetroffene und marginalisierte Menschen, die ohnehin mit strukturellen Barrieren in der Medizin konfrontiert sind. Um eine algorithmische Zwei-Klassen-Medizin und den kontrollierten Abfluss von Gesundheitsdaten zu verhindern, fordern wir lückenlose Transparenz.

In diesem Zusammenhang bitten wir den Stadtrat um die Beantwortung der folgenden Fragen:

1. Welche konkreten KI-Systeme und Softwareprodukte (Produktname und Hersteller/Anbieter) sind derzeit in den Dienstabteilungen des GUD und des Stadtspitals im Einsatz (bitte nach Einsatzbereich, klinischer bzw. administrativer Funktion und Anbieter aufschlüsseln)?
2. Welche weiteren spezifischen KI-Systeme oder -Anwendungen befinden sich aktuell in der Planung, in der Evaluation oder in einer Pilotphase?
3. Werden im Rahmen der aktuellen oder geplanten KI-Anwendungen Patient*innendaten oder andere Gesundheitsdaten direkt oder indirekt für das Training, das Fine-Tuning oder die allgemeine Weiterentwicklung von Modellen verwendet?



2/10

4. Falls die Verwendung von Gesundheitsdaten für das Training oder Fine-Tuning zutrifft (siehe Frage 3): Erfolgt diese Datenverarbeitung und das Modelltraining ausschliesslich innerhalb der städtischen IT-Infrastruktur (OIZ) oder auch auf Systemen externer Anbieter?
5. Kann der Stadtrat vertraglich und technisch absolut ausschliessen, dass sensible Daten aus dem Stadtzürcher Gesundheitswesen in kommerzielle Modelle ausländischer Anbieter einfliessen und dort ausserhalb der Kontrolle der Stadt weiterverwendet werden?
6. Auf welchen Datensätzen basiert die im GUD eingesetzten Systeme primär (z.B. internationale, herstellerspezifische Trainingsdaten oder institutionsspezifische Daten aus der Schweiz)?
7. Welche der im GUD und im Stadtspital eingesetzten oder geplanten medizinischen KI-Anwendungen stammen entweder direkt von US-Technologiekonzernen (z.B. Microsoft, Google, AWS, Palantir, Oracle) oder nutzen deren KI-Modelle und/oder Infrastrukturen im Hintergrund als integrierte White-Label oder Schnittstellen-Lösungen (APIs)?
8. Wie stellt der Stadtrat sicher, dass US-Behörden über den US CLOUD Act oder FISA 702 Zugriff auf städtische Patient*innendaten erlangen können, wenn die Verträge zwar mit Schweizer Tochtergesellschaften abgeschlossen wurden, die technologische Infrastruktur (Server, Backup-Systeme, Hosting) jedoch von US-Mutterkonzernen kontrolliert wird?
9. Werden die in die Cloud übertragenen Patient*innendaten ausschliesslich im Ruhezustand (Data at rest) und bei der Übertragung (Data in transit) verschlüsselt, oder kommt durchgehend Confidential Computing (Verschlüsselung auch während der aktiven Verarbeitung im Arbeitsspeicher der Cloud-Server) zum Einsatz, um ein Mitlesen durch den Cloud-Anbieter oder ausländische Geheimdienste technisch unmöglich zu machen?
10. Schliessen die Verträge mit den Software-Anbietern explizit aus, dass Metadaten, Telemetriedaten oder anonymisierte/pseudonymisierte klinische Verlaufsdaten zur algorithmischen Optimierung, zum «Alignment» oder zum allgemeinen Training kommerzieller, herstellereigener Modelle ausserhalb der Schweiz verwendet werden? Wo wird dies unabhängig auditiert?
11. Wie stellt der Stadtrat sicher, dass die diagnostischen KI-Systeme (insbesondere in der Radiologie, Gastroenterologie und Dermatologie) auf Bild- und Trainingsdatensätzen basieren, die die anatomischen, physiologischen, dermatologischen und symptomatischen Spezifika von BIPOC (Black, Indigenous, People of Color) und FINTA (Frauen, intergeschlechtliche, nicht-binäre, trans- und agender Personen) adäquat abbilden, um KI-bedingte Fehldiagnosen zu verhindern?
12. Werden die Fehlerraten und die diagnostische Präzision der im Stadtspital eingesetzten KI-Systeme systematisch differenziert ausgewertet, um zu verhindern, dass beispielsweise bei FINTA (Frauen, intergeschlechtliche, nicht-binäre, trans- und agender Personen) oder Women of Color schlechtere medizinische Ergebnisse erzielt werden als bei der weissen, männlichen Referenzgruppe?
13. Wie wird verhindert, dass bei einer allfälligen Anwendung von, prädiktiven KI-Systemen (z.B. durch PROMs oder bei administrativen Triage-Entscheidungen) sozioökonomische Faktoren wie Armut, prekäre Wohnverhältnisse, Bildungsstand oder das Vorliegen einer Zusatzversicherung als negative Indikatoren bewertet werden und dadurch armutsbetroffenen Menschen den Zugang zu adäquaten Therapien erschweren oder verweigern?
14. Werden Patient*innen im Stadtspital aktiv und verständlich darüber aufgeklärt, wenn eine Diagnose, eine Operationsplanung oder eine therapeutische Empfehlung massgeblich von einer KI generiert oder beeinflusst wurde? Gibt es eine transparente Kennzeichnungspflicht in der Patient*innenakte?

Der Stadtrat beantwortet die Anfrage wie folgt:

Im Stadtspital Zürich (STZ) sind derzeit nur wenige KI-gestützte Lösungen im Einsatz, die ausschliesslich in unterstützender Funktion angewendet werden. Es handelt sich nicht um eigen-



3/10

ständig entscheidende, sondern um assistierende Systeme, deren Ergebnisse durch Fachpersonal überprüft und verantwortet werden. In den Gesundheitszentren für das Alter (GFA), den städtischen Gesundheitsdiensten (SGD) und im Umwelt- und Gesundheitsschutz (UGZ) sind aktuell keine produktiven KI-Systeme im klinischen, pflegerischen oder administrativen Kernbetrieb im Einsatz, wobei gelegentlich generative KI-Services als unterstützende Arbeitsmittel genutzt werden.

Für das STZ ist der aktive KI-Einsatz eine medizinische, prozessuale und wirtschaftliche Notwendigkeit. Das STZ treibt daher den gezielten und verantwortungsvollen KI-Einsatz aktiv voran, mit klarem Fokus auf gleichberechtigte Versorgung aller Patientinnen und Patienten, Datenschutz und Transparenz. Geplante Pilotversuche durchlaufen den städtischen ISDS-Prozess. In diesem Rahmen erfolgt eine Involvierung der Fachstelle Informationssicherheit der Organisation und Informatik Zürich (OIZ) und der Datenschutzstelle, falls eine beabsichtigte Bearbeitung von Personendaten mit besonderen Risiken für die Grundrechte der Betroffenen verbunden ist.

Nach diesen einleitenden Bemerkungen können die Fragen wie folgt beantwortet werden:

Frage 1

Welche konkreten KI-Systeme und Softwareprodukte (Produktname und Hersteller/Anbieter) sind derzeit in den Dienstabteilungen des GUD und des Stadthospitals im Einsatz (bitte nach Einsatzbereich, klinischer bzw. administrativer Funktion und Anbieter aufschlüsseln)?

Im STZ sind derzeit nur wenige KI-gestützte Lösungen und ausschliesslich in unterstützender Funktion im Einsatz. Es handelt sich nicht um eigenständig entscheidende, sondern um assistierende Systeme, deren Ergebnisse durch Fachpersonal überprüft und verantwortet werden. Folgende Produkte stehen im Einsatz:

Produkt	Anbieter	Einsatzbereich	Funktion
RapidAI	RapidAI	Bildgebung (MRT, CT)	klinisch
ERAS Interactive Audit System (EIAS)	Encare	Vor-/Nachsorge bei Operationen	administrativ
VisiblePatient	Visible Patient	OP-Vorbereitung (zur Verfügung stellen von 3D-Modellen von Organen)	klinisch
Bitfount	Bitfount	Klinische Forschung	administrativ
ZüriA	OIZ	Supportbereiche	administrativ

In den GFA und SGD sind aktuell keine produktiven KI-Systeme im klinischen, pflegerischen oder administrativen Kernbetrieb im Einsatz. Gelegentlich werden generative KI-Services wie ZüriA, Microsoft Cockpit und ChatGPT als unterstützende Arbeitsmittel für Textarbeit und Strukturierung genutzt, stets unter Einhaltung der städtischen Vorgaben und ausschliesslich mit nicht vertraulichen Inhalten.

Im UGZ sind folgende KI-Systeme im Einsatz:



4/10

Produkt	Anbieter	Einsatzbereich	Funktion
Züri-Bot	Abraxas	Digitale Erstberatung zu Heizungsersatz und energetischer Gebäudesanierung.	Beratung
KITRO	KITRO	Food-Waste-Messsystem: KI-gestützte Erfassung und Auswertung von Lebensmittelabfällen im Bereich Ernährung	Analyse

Frage 2

Welche weiteren spezifischen KI-Systeme oder -Anwendungen befinden sich aktuell in der Planung, in der Evaluation oder in einer Pilotphase?

Künstliche Intelligenz verändert Medizin und Spitalbetrieb grundlegend und der Einsatz von Produkten, die auf entsprechenden Technologien basieren, ist auch im Gesundheitswesen zunehmend unverzichtbar. Für das STZ ist deshalb der aktive KI-Einsatz keine Option, sondern eine medizinische, prozessuale und wirtschaftliche Notwendigkeit. Spitäler, die hierin nicht investieren, fallen in Qualität, Effizienz und Attraktivität zurück. Das STZ treibt daher den gezielten und verantwortungsvollen KI-Einsatz aktiv voran, mit klarem Fokus auf gleichberechtigte Versorgung aller Patientinnen und Patienten, Datenschutz und Transparenz.

- Dragon Co-Pilot und KISIM AI: Für die automatisierte Dokumentation auf Basis von Gesprächstranskriptionen oder Diktaten (Ambient Listening, Berichtbeschreibung).
- Cody: Für die medizinische Codierung.

Beide Vorhaben befinden sich noch in der Vorprüfung. Das Ziel der Pilotversuche ist es, derartige Lösungen unter kontrollierten Bedingungen zu evaluieren und ihre Vor- und allfälligen Nachteile zu ermitteln. In beiden Bereichen – Transkription und Codierung – besteht von vornherein kein Risiko einer klinischen Diskriminierung oder eines für Minderheiten nachteiligen Bias, da die Lösungen weder diagnostische noch prädiktive Bewertungen von Personen vornehmen.

Ebenfalls in Prüfung ist das Projekt Digital Eye Clinic (Eigenentwicklung der Augenklinik in Zusammenarbeit mit Forschungspartnern) sowie eine Ausschreibung für eine digitale Sturzprävention mit KI-gestützter Bewegungsanalyse. Das Risiko eines diskriminierenden Bias existiert auch hier von vornherein nicht.

Für die erwähnten Vorhaben muss der städtische ISDS-Prozess durchlaufen werden.

Schliesslich werden verschiedene Anwendungen im Bereich der Radiologie geprüft. Es geht um die teilautomatisierte Unterstützung bei der Auswertung von radiologischen Bilddaten.

In den GFA werden zwei Vorhaben vorbereitet:

- Ein **sozialer Roboter von Navel Robotics** zur sozialen Interaktion und Aktivierung, mit lokaler Datenverarbeitung und ohne Cloud-Anbindung. Der Pilot-Einsatz ist voraussichtlich ab September 2026 in einem dedizierten Bereich eines Betriebs geplant und erfolgt vorbehältlich der erforderlichen Freigaben aus dem laufenden städtischen ISDS-Prozess.



5/10

- Eine mögliche **KI-gestützte Unterstützung der Dienstplanung** in Zusammenarbeit mit Nexus Schweiz AG. Das Vorhaben befindet sich in einer frühen Abklärungsphase und betrifft Mitarbeitende, nicht Patient*innen.

In den SGD befinden sich keine spezifische KI-Systeme in Planung oder Evaluation.

Im UGZ läuft der Pilot Zü-Re, ein öffentlich zugänglicher KI-Chatbot zum Teilen, Weitergeben und Reparieren (Kreislaufwirtschaft). Zü-Re greift für die Antworten auf eine vom UGZ kuratierte Datensammlung zu. Das Pilotprojekt läuft bis Ende 2026.

Frage 3

Werden im Rahmen der aktuellen oder geplanten KI-Anwendungen Patient*innendaten oder andere Gesundheitsdaten direkt oder indirekt für das Training, das Fine-Tuning oder die allgemeine Weiterentwicklung von Modellen verwendet?

Die im STZ eingesetzten unterstützenden Lösungen sowie die geplanten Pilotversuche verwenden in der Regel keine Daten von Patientinnen oder Patienten oder sonstigen Gesundheitsdaten des STZ für das Training, das Fine-Tuning oder die allgemeine Weiterentwicklung von Modellen der Anbietenden. Eine entsprechende Sekundärnutzung wird vertraglich ausgeschlossen und soweit möglich technisch unterbunden. Sofern weder vertraglicher Ausschluss oder noch technische Unterbindung möglich sind, werden Patientinnen-/Patienten- oder sonstige Gesundheitsdaten ausschliesslich in vorweg gegenüber dem Anbieter de-identifizierter Form verwendet und das nur mit dem Einverständnis der Patientinnen oder Patienten. Die Einhaltung der geltenden datenschutzrechtlichen Rahmenbedingungen bildet einen festen Bestandteil der datenschutzrechtlichen Prüfung vor der Einführung. Dabei ist festzuhalten, dass die anwendbare Datenschutzgesetzgebung eine solche Nutzung von personenbezogenen Daten für nicht personenbezogene Zwecke nicht von vornherein ausschliesst und sie unter gewissen Voraussetzungen ermöglicht. Es ist somit nicht auszuschliessen, dass das STZ diesbezüglich in Einzelfällen Ausnahmen gewährt. Solche Ausnahmen würden allerdings voraussetzen, dass das STZ nach eingehender Prüfung zum Schluss kommt, dass die Vorteile (insbesondere für die Patientinnen und Patienten des STZ) stärker wiegen als die Nachteile. Zudem müssen solche Vorhaben den städtischen ISDS-Prozess durchlaufen.

Nach aktuellem Stand werden bei den GFA keine Patientinnen-/Patienten-, Bewohnenden- oder Gesundheitsdaten für das Training, Fine-Tuning oder die allgemeine Weiterentwicklung von KI-Modellen verwendet. Das gilt auch für die beiden derzeit geprüften oder vorbereiteten Vorhaben. Beim Vorhaben zur Dienstplanung betrifft die Abklärung nach aktuellem Verständnis die Mitarbeitenden- und Planungsdaten.

Frage 4

Falls die Verwendung von Gesundheitsdaten für das Training oder Fine-Tuning zutrifft (siehe Frage 3): Erfolgt diese Datenverarbeitung und das Modelltraining ausschliesslich innerhalb der städtischen IT-Infrastruktur (OIZ) oder auch auf Systemen externer Anbieter?

Da in den Dienstabteilungen des GUD keine Gesundheitsdaten für das Training oder Fine-Tuning verwendet werden, stellt sich die Frage nach dem Ort der Datenverarbeitung nicht.



6/10

Frage 5

Kann der Stadtrat vertraglich und technisch absolut ausschliessen, dass sensible Daten aus dem Städtzürcher Gesundheitswesen in kommerzielle Modelle ausländischer Anbieter einfließen und dort ausserhalb der Kontrolle der Stadt weiterverwendet werden?

Eine absolute, jede denkbare Möglichkeit ausschliessende Garantie kann bei keiner informationstechnischen Verarbeitung redlich abgegeben werden. Das STZ und die anderen Dienstabteilungen des GUD verhindern jedoch wirksam einen unzulässigen Abfluss von Daten durch:

- vertragliche Ausschlüsse jeglicher Sekundärnutzung von Gesundheitsdaten durch Anbietende
- technische Schutzmassnahmen wie Datenminimierung und De-Identifizierung

Im Rahmen des städtischen ISDS-Prozesses werden die von den Dienstabteilungen vorgesehenen Sicherheitsmassnahmen geprüft.

Frage 6

Auf welchen Datensätzen basiert die im GUD eingesetzten Systeme primär (z.B. internationale, herstellerepezifische Trainingsdaten oder institutionsspezifische Daten aus der Schweiz)?

Die im Stadtsptal eingesetzten unterstützenden Lösungen beruhen auf den vom jeweiligen Hersteller bereitgestellten und nach anerkannten wissenschaftlichen Massstäben entwickelten Datengrundlagen. Eine institutionsspezifische Weiterentwicklung anhand von Daten des STZ findet nicht statt. Die konkrete Zusammensetzung der herstellereitigen Trainingsdaten liegt im Verantwortungsbereich der Anbietenden und ist Gegenstand der Anbietendenbewertung im Rahmen der Einführungsprüfung. In der Regel dürften KI-gestützte Lösungen, die aktiv für medizinische Zwecke (Diagnose, Behandlungsempfehlungen / -entscheide, etc.) eingesetzt werden, als Medizinprodukte qualifizieren und Gegenstand der entsprechenden regulatorischen Anforderungen sein.

In den anderen Dienstabteilungen (GFA, SGD, UGZ) werden keine KI-Systeme eingesetzt, die auf institutionsspezifischen Trainingsdaten basieren. Die verwendeten Tools (z. B. ZüriA, ZüRe) nutzen allgemeine Sprachmodelle, deren Trainingsdaten nicht aus dem Städtzürcher Gesundheitswesen stammen.

Frage 7

Welche der im GUD und im Stadtsptal eingesetzten oder geplanten medizinischen KI-Anwendungen stammen entweder direkt von US-Technologiekonzernen (z.B. Microsoft, Google, AWS, Palantir, Oracle) oder nutzen deren KI-Modelle und/oder Infrastrukturen im Hintergrund als integrierte White-Label oder Schnittstellen-Lösungen (APIs)?

Im STZ nutzen die geplanten Pilotanwendungen Dragon Co-Pilot und Cody Infrastrukturen von Microsoft (Azure EU, bzw. Azure Switzerland).



7/10

Soweit eingesetzte oder geplante Lösungen auf Modellen oder Infrastrukturen entsprechender Anbietender aufbauen, müssen diese den städtischen ISDS-Prozess durchlaufen. Sie unterliegen den unter den Fragen 5, 8, 9 und 10 beschriebenen vertraglichen und technischen Schutzmassnahmen. Keine der eingesetzten oder geplanten medizinischen KI-Anwendungen stammt direkt von Palantir, Google, AWS oder Oracle. Die verwendeten Systeme sind spezialisierte Medizinprodukte, nicht allgemeine Cloud-KI-Plattformen.

In den anderen Dienstabteilungen (GFA, SGD, UGZ) werden keine medizinischen KI-Anwendungen eingesetzt. Die genutzten Cloud-KI-Tools wie ChatGPT nutzen zwar Infrastrukturen von US-Unternehmen, jedoch nicht für medizinische Zwecke und nicht mit sensiblen Daten. ZüriA läuft auf der Infrastruktur der OIZ.

Frage 8

Wie stellt der Stadtrat sicher, dass US-Behörden über den US CLOUD Act oder FISA 702 Zugriff auf städtische Patient*innendaten erlangen können, wenn die Verträge zwar mit Schweizer Tochtergesellschaften abgeschlossen wurden, die technologische Infrastruktur (Server, Backup-Systeme, Hosting) jedoch von US-Mutterkonzernen kontrolliert wird?

Bei sämtlichen Projekten und Vorhaben der Stadtverwaltung, bei denen Informationen und Daten bearbeitet werden, muss ein ISDS-Prozess durchlaufen werden. Im ISDS-Prozess wird systematisch geprüft, ob ein geplantes IT-Projekt oder eine Datenverarbeitung den Sicherheits- und Datenschutzanforderungen der Stadt Zürich entspricht. Dabei werden die rechtlichen Grundlagen, das Risiko für die Vertraulichkeit, die Integrität und Verfügbarkeit der Daten, die technische Architektur, die eingesetzten Schutz- und Verschlüsselungsmassnahmen sowie die Verträge mit externen Dienstleistenden geprüft. Bezogen auf die konkreten Risiken werden geeignete technische und organisatorische Massnahmen ergriffen, um die Einhaltung der Sicherheits- und Datenschutzanforderungen der Stadt Zürich, die auch die Verhinderung eines Zugriffs durch unbefugte Personen umfasst, zu gewährleisten.

Eine mögliche technische Massnahme zur Reduktion des Risikos von Zugriffen ist die Verschlüsselung von Daten. In organisatorischer Hinsicht ist unter anderem die vertragliche Vereinbarung von Geheimhaltungspflichten gemäss § 25 der Verordnung über die Information und den Datenschutz (IDV, LS 170.41) eine geeignete Massnahme dafür.

Das Stadtspital begegnet dem Risiko eines extraterritorialen behördlichen Zugriffs durch Datenminimierung, durch technische Schutzmassnahmen, die ein Mitlesen durch den Anbietenden verhindern (vgl. Antwort zu Frage 9), durch vertragliche Vorgaben zu Datenstandort und Sekundärnutzung. Soweit ein Anbieter mittelbar einer ausländischen Konzernmutter zugeordnet ist, wird diesem Umstand im Rahmen der Einführungsprüfung Rechnung getragen; massgebend ist, dass ein Zugriff auf Klartextdaten durch technische und organisatorische Massnahmen ausgeschlossen, beziehungsweise soweit möglich technisch unmöglich gemacht wird.

In den anderen Dienstabteilungen (GFA, SGD, UGZ) werden keine Patientinnen- oder Patientendaten in die Cloud übertragen, weshalb das Risiko zurzeit nicht besteht. Neue KI-Tools würden denselben Regeln und Prinzipien unterliegen wie oben beschrieben.



8/10

Frage 9

Werden die in die Cloud übertragenen Patient*innendaten ausschliesslich im Ruhezustand (Data at rest) und bei der Übertragung (Data in transit) verschlüsselt, oder kommt durchgehend Confidential Computing (Verschlüsselung auch während der aktiven Verarbeitung im Arbeitsspeicher der Cloud-Server) zum Einsatz, um ein Mitlesen durch den Cloud-Anbieter oder ausländische Geheimdienste technisch unmöglich zu machen?

Daten von Patientinnen und Patienten werden im STZ im Ruhezustand (data at rest) und bei der Übertragung (data in transit) verschlüsselt. Für Verarbeitungen, bei denen ein Mitlesen durch den Anbietenden ausgeschlossen werden soll, wird der Einsatz von Confidential Computing (Verschlüsselung auch während der aktiven Verarbeitung in hardwarebasierten vertrauenswürdigen Ausführungsumgebungen) geprüft oder als Schutzmassnahme vorgesehen, wo das technisch verfügbar und angezeigt ist. Die konkrete Schutzarchitektur wird je Lösung im Rahmen der Einführungsprüfung festgelegt.

In den anderen Dienstabteilungen (GFA, SGD, UGZ) werden keine Patientinnen- oder Patientendaten in die Cloud übertragen, weshalb die Frage nicht anwendbar ist. Auch beim Pilotprojekt der GFA mit dem sozialen Roboter Navel (der jedoch keine Patientinnen- und Patientendaten verwendet) werden die Daten lokal auf dem Gerät verbleiben und verschlüsselt gespeichert werden.

Frage 10

Schliessen die Verträge mit den Software-Anbietern explizit aus, dass Metadaten, Telemetriedaten oder anonymisierte/pseudonymisierte klinische Verlaufsdaten zur algorithmischen Optimierung, zum «Alignment» oder zum allgemeinen Training kommerzieller, herstellereigener Modelle ausserhalb der Schweiz verwendet werden? Wo wird dies unabhängig auditiert?

Ja, die Verträge mit den Anbietenden schliessen die Nutzung von Gesundheitsdaten – einschliesslich Metadaten, Telemetriedaten sowie anonymisierter oder pseudonymisierter klinischer Verlaufsdaten – zur algorithmischen Optimierung, zum «Alignment» oder zum allgemeinen Training herstellereigener Modelle ausserhalb des vereinbarten Zwecks aus.

Die Einhaltung wird gesichert durch:

- Anbieterzertifizierungen (z. B. ISO 27001, SOC 2)
- vertraglich vereinbarte Audit- und Kontrollrechte

Wie bereits ausgeführt, unterliegen solche Vorhaben auch dem städtischen ISDS-Prozess.

In den anderen Dienstabteilungen (GFA, SGD, UGZ) werden keine klinischen Verlaufsdaten verarbeitet.



9/10

Frage 11

Wie stellt der Stadtrat sicher, dass die diagnostischen KI-Systeme (insbesondere in der Radiologie, Gastroenterologie und Dermatologie) auf Bild- und Trainingsdatensätzen basieren, die die anatomischen, physiologischen, dermatologischen und symptomatischen Spezifika von BIPOC (Black, Indigenous, People of Color) und FINTA (Frauen, intergeschlechtliche, nicht-binäre, trans- und agender Personen) adäquat abbilden, um KI-bedingte Fehldiagnosen zu verhindern?

Die Dienstabteilungen des GUD setzen in den genannten Bereichen (Radiologie, Gastroenterologie, Dermatologie) derzeit keine eigenständig diagnostizierenden KI-Systeme in der Versorgung ein; die vorhandenen Lösungen wirken unterstützend, die Diagnose verantwortet das Fachpersonal.

Der Stadtrat stellt fest, dass mögliche Verzerrungen aufgrund fehlender Datensätze von Referenzgruppen sich auf der Ebene des Spitals als Nutzende einer KI-gestützten Lösung nicht wirksam beheben lassen. KI-gestützte medizinische Lösungen bauen – wie auch ärztliches Handeln – auf dem anerkannten wissenschaftlichen Erkenntnisstand auf. Wo der Erkenntnisstand selbst Verzerrungen aufweist, ist dem auf der Ebene der wissenschaftlichen und herstellerseitigen Grundlagen zu begegnen. KI kann solche bereits in den Grundlagen angelegten Verzerrungen verstärken; die Ursache liegt jedoch nicht in der KI als Werkzeug, sondern in den ihr zugrunde liegenden wissenschaftlichen Daten.

Das STZ trägt diesem Umstand Rechnung, indem es bei der Einführung jeder Lösung die im Einzelfall bestehenden Risiken ermittelt, die Aussagekraft der herstellerseitigen Datengrundlagen bewertet und geeignete Massnahmen vorsieht, einschliesslich der durchgängigen fachlichen Überprüfung der Ergebnisse. Die Verantwortung für Behandlungsentscheidungen wird wie bereits erwähnt auch dann bei den behandelnden Ärztinnen und Ärzten verbleiben, wenn das STZ in Zukunft stärker auf KI-gestützte medizinische Anwendungen setzt.

Frage 12

Werden die Fehlerraten und die diagnostische Präzision der im Stadtspital eingesetzten KI-Systeme systematisch differenziert ausgewertet, um zu verhindern, dass beispielsweise bei FINTA (Frauen, intergeschlechtliche, nicht-binäre, trans- und agender Personen) oder Women of Color schlechtere medizinische Ergebnisse erzielt werden als bei der weissen, männlichen Referenzgruppe?

Da derzeit keine eigenständig diagnostizierenden KI-Systeme in der Versorgung eingesetzt werden, findet im STZ eine entsprechende differenzierte Auswertung gegenwärtig nicht statt. Für einen allfälligen künftigen Einsatz solcher Systeme wird die Frage einer nach relevanten Untergruppen differenzierten Leistungs- und Fehlerbewertung Teil der einzelfallbezogenen Risikobeurteilung sein. Es gelten sinngemäss die Ausführungen zu Frage 11. Die fachliche Letztverantwortung verbleibt beim Personal.



10/10

Frage 13

Wie wird verhindert, dass bei einer allfälligen Anwendung von, prädiktiven KI-Systemen (z.B. durch PROMs oder bei administrativen Triage-Entscheidungen) sozioökonomische Faktoren wie Armut, prekäre Wohnverhältnisse, Bildungsstand oder das Vorliegen einer Zusatzversicherung als negative Indikatoren bewertet werden und dadurch armutsbetroffenen Menschen den Zugang zu adäquaten Therapien erschweren oder verweigern?

STZ, SGD und GFA gewährleisten Zugang zu qualitativ hochwertigen medizinischen Leistungen unabhängig von Herkunft, Versicherungsstatus, Geschlecht, Sprache, Religion oder Kultur der Patientinnen und Patienten.

Die Dienstabteilungen des GUD setzen derzeit keine KI-Systeme ein, die sozioökonomische Faktoren wie Armut, Wohnverhältnisse, Bildungsstand oder das Vorliegen einer Zusatzversicherung als Indikatoren in Triage- oder Therapieentscheidungen einfließen lassen. Sowohl die bereits eingesetzten Tools (Antwort zur Frage 1) als auch die Pilotvorhaben (Antwort zur Frage 2) nehmen keine prädiktiven Bewertungen von Personen vor.

Frage 14

Werden Patient*innen im Stadtspital aktiv und verständlich darüber aufgeklärt, wenn eine Diagnose, eine Operationsplanung oder eine therapeutische Empfehlung massgeblich von einer KI generiert oder beeinflusst wurde? Gibt es eine transparente Kennzeichnungspflicht in der Patient*innenakte?

Die derzeit eingesetzten Lösungen wirken unterstützend; die ärztliche Verantwortung für Diagnose, Operationsplanung und therapeutische Empfehlungen verbleibt beim Fachpersonal. Über den Einsatz von Datenverarbeitungslösungen werden Patientinnen oder Patienten im Rahmen der allgemeinen Datenschutzinformation und, wo angezeigt, durch spezifische Transparenzhinweise informiert; entsprechende Hinweise sind fester Bestandteil der für die Pilotvorhaben erstellten Datenschutzkonzepte.

Die Frage einer formellen Kennzeichnungspflicht in der Patientendokumentation für massgeblich KI-beeinflusste Entscheide wird im Zuge der Weiterentwicklung des KI-Einsatzes beurteilt.

Im Namen des Stadtrats
Der Stadtschreiber
Thomas Bolleter