Stadt Zürich
Gemeinderat

GR Nr. 2026/ 282

10. Juni 2026

Schriftliche Anfrage

von Vera Çelik (SP)
Yves Henz (Grüne)

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 Insellösungen sind und massive Investitionskosten verursachen. Auch die rechtliche Haftung und der Datenschutzsind 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?
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?