Hamburger Thesen zum Personenbezug von Large Language Models

Thomas Fuchs*

Seit der Veröffentlichung von ChatGPT im November 2022 ist „KI“, genauer der Einsatz von Large Language Models (LLMs) in KI-Systemen, ein großes Thema in Wirtschaft und Gesellschaft. Vor dem Hintergrund der rasanten Entwicklung und der explodierenden Breite der Anwendungen fällt auf, dass es bislang kaum Festlegungen von Datenschutzaufsichtsbehörden zu diesem Komplex gibt. Dies ist fraglos auf die Komplexität der neuen Technologien zurückzuführen, auf die Herausforderung, geltendes Recht wie die DSGVO auf diese sich schnell fortentwickelnden Technologien anzuwenden, und sicher auch auf die Sorge, Konsequenzen nicht überblicken oder verantworten zu können.

So ist es auch bei der grundlegenden Frage, ob LLMs überhaupt personenbezogene Daten im Sinne der DSGVO speichern. Mit den folgenden, am 15. 7. 2024 veröffentlichten Hamburger Thesen möchte der HmbBfDI nun einen Impuls für die Debatte um die datenschutzrechtliche Einordnung von LLMs setzen:

1. Die bloße Speicherung eines LLMs stellt keine Verarbeitung im Sinne des Art. 4 Nr. 2 DSGVO dar. Denn in LLMs werden keine personenbezogenen Daten gespeichert. Soweit in einem LLM-gestützten KI-System personenbezogene Daten verarbeitet werden, müssen die Verarbeitungsvorgänge den Anforderungen der DSGVO entsprechen. Dies gilt insbesondere für den Output eines solchen KI-Systems.
2. Mangels Speicherung personenbezogener Daten im LLM können die Betroffenenrechte der DSGVO nicht das Modell selbst zum Gegenstand haben. Ansprüche auf Auskunft, Löschung oder Berichtigung können sich jedoch zumindest auf Input und Output eines KI-Systems beziehen.
3. Das Training von LLMs mit personenbezogenen Daten muss datenschutzkonform erfolgen. Dabei sind auch die Betroffenenrechte zu beachten. Ein ggf. datenschutzwidriges Training wirkt sich aber nicht auf die Rechtmäßigkeit des Einsatzes eines solchen Modells in einem KI-System aus.

Grundlage dieser Thesen ist die wichtige Unterscheidung zwischen KI-Systemen und den darin enthaltenen LLMs. Ein KI-System (etwa: ChatGPT) besteht aus mehreren Komponenten, von denen das LLM (etwa: GPT 4 o) nur eine darstellt. Zu den weiteren Bausteinen zählen etwa die Benutzerschnittstelle sowie Ein- und Ausgangsfilter sowie ggf. Anreicherungsprozesse, z. B. durch Datenbankabfragen, Internetsuchen oder RAG.

Die Funktionsweise von LLMs unterscheidet sich fundamental von herkömmlichen Datenspeichermethoden. LLMs operieren mit abstrakten mathematischen Repräsentationen. Trainingstexte werden in numerische Tokens zerlegt – kleinere Einheiten als Wörter, aber größer als einzelne Buchstaben. Im Trainingsprozess lernt das LLM, die Beziehungen zwischen diesen Tokens zu verstehen und die Wahrscheinlichkeiten für bestimmte Wortfolgen einzuschätzen, hieraus bilden sich Beziehungsverflechtungen dieser Token, sog. Embeddings. Dieses erlernte Wissen bildet die Grundlage für die Fähigkeit des LLMs, kohärente Texte zu erzeugen.

Soweit das Trainingsmaterial personenbezogene Daten enthält, durchlaufen diese eine Transformation; die konkreten Bezüge zu bestimmten Personen gehen verloren. Stattdessen werden allgemeine Muster und Zusammenhänge extrahiert, die sich aus der Gesamtheit der Trainingsdaten ergeben. Was LLMs produzieren, ist damit nicht die Wiedergabe gespeicherter Informationen. Im Gegensatz zu vom EuGH anerkannten personenbezogenen Identifiern wie IP-Adressen dienen die gespeicherten Tokens und Embeddings nicht der gezielten, personenindividuellen Zuordnung. Dieser Ansatz bedeutet keine Schwächung des Datenschutzes, sondern im Gegenteil eine notwendige Präzisierung der Verantwortlichkeiten. Ziel ist es, den Datenschutz dort zu stärken, wo er am effektivsten wirken kann: bei der Nutzung von KI-Systemen, also beim Output, und bei dem Training der Modelle. Der Ansatz erkennt die real bestehenden Unterschiede zwischen einer Datenbank und einem LLM an und kommt zudem zu praktikablen Ergebnissen.

Die Thesen sind bewusst als Diskussionspapier formuliert. In einer sich kontinuierlich verändernden digitalen Welt reicht es nicht, nach mehrjähriger Prüfung vermeintlich abschließende Positionspapiere zu veröffentlichen. Wir müssen uns auch mit Thesen der Diskussion stellen – auch auf die Gefahr hin, sich angreifbar oder widerlegbar zu machen. Gleichzeitig bleiben wir offen für die Weiterentwicklung der Thesen durch Rückmeldungen aus Praxis und Wissenschaft. Gemeinsam mit allen Beteiligten wollen wir den Weg zu einem zukunftsfähigen Datenschutzrecht ebnen, das Innovationen ermöglicht und gleichzeitig die Rechte und Freiheiten der Bürger:innen im Kontext fortschreitender KI-Entwicklungen wirksam schützt.

* Thomas Fuchs, LL.M. Eur., ist seit 2021 Hamburgischer Beauftragter für Datenschutz und Informationsfreiheit. Zuvor war er von 2008 bis 2021 Direktor der Medienanstalt Hamburg/Schleswig-Holstein.