

Datenschutzrechtliche Analyse: Ollama

Stand: September 2025

I. Zu bewertendes Verfahren / Tool und Zweckbestimmung

Ollama ist ein Open-Source-Tool, das entwickelt wurde, um das lokale Ausführen und Verwalten großer Sprachmodelle (LLMs) auf dem eigenen Computer zu erleichtern. Während LLM-Modelle wie GPT-4 von Open AI in sich geschlossen und damit in datenschutzrechtlich bedenklicher Weise intransparent sind, bietet Ollama eine Open-Source-Plattform, die eine einfache Möglichkeit bietet, Modelle wie z.B. Llama, Mistral oder Gemma herunterzuladen, zu konfigurieren und lokal auszuführen. Dadurch bleiben alle Daten im Unternehmen, und werden nicht an externe Cloud-Anbieter übermittelt.

1. Betroffenengruppen deren personenbezogene Daten verarbeitet werden

Ollama ist lediglich ein Tool, das eine reibungslose Installation eines LLM wie z.B. Mistral auf dem eigenen Server auf einfachem Weg ermöglichen soll. Das Tool selbst verarbeitet keine personenbezogenen Daten. Eine Datenschutzfolgeabschätzung wäre an diesem Punkt zu Ende, denn von Ollama geht kein Datenschutzrisiko aus.

Angesichts dessen, dass die Implementierung von LLMs momentan ein bedeutendes Thema ist, stellt sich die Frage, inwieweit der Einsatz von Ollama aus datenschutzrechtlicher Sicht hilfreich dabei sein kann.

Das LLM selbst (ob Chat GPT oder Mistral) kann keine Unterscheidung zwischen personenbezogenen und nicht personenbezogenen Daten treffen und auch eine dahingehende technische Modifikation durch Filter ist bislang nur sehr eingeschränkt möglich.

Die sich daraus ergebenden datenschutzrechtlichen Herausforderungen sind unter anderem Folgende:

1) Verarbeitung personenbezogener Daten

LLMs werden mit großen Mengen an Textdaten trainiert. Einige LLM nutzen auch die beim prompten eingegebenen Daten der User als Trainingsdaten zur Weiterentwicklung ihrer KI. Eine Anweisung an die Mitarbeitenden, beim prompten darauf zu achten keine personenbezogenen Daten einzugeben, ist hilfreich, beinhaltet aber auch das Risiko, dass sich nicht alle Mitarbeitenden zu jeder Zeit daranhalten.

Gefördert vom

Analyse gem. RL zur Risikoanalyse und Risikobehandlung im Rahmen des KDM - Seite 2

Herausforderung: Die DS-GVO verlangt eine rechtmäßige Grundlage (z. B. Einwilligung, Vertragserfüllung, berechtigtes Interesse) für die Verarbeitung personenbezogener Daten. Bei LLMs ist oft unklar, ob und wie diese Grundlage für Trainingsdaten vorliegt. Für die Verarbeitung der Inhalte von Prompts gibt es, wenn es sich um personenbezogene Daten handelt, keine Rechtsgrundlage. Die Verarbeitung ohne Rechtsgrundlage ist ein Datenschutzverstoß, der mit Bußgeldern sanktioniert wird.

2) Transparenz und Betroffenenrechte

LLMs sind oft nicht transparent. User wissen nicht, welche Daten verarbeitet wurden. Die DS-GVO verlangt jedoch Auskunftsrechte (Art. 15) und das Recht auf Löschung (Art. 17).

Herausforderung: Wie können Betroffene ihre Rechte geltend machen, wenn das Modell keine klare Datenherkunft oder -speicherung hat?

3) Datenminimierung und Speicherbegrenzung

LLMs speichern oft unbestimmt lange Daten aus Eingaben (z. B. Chatverläufe). Art. 5 Abs. 1 lit. c und e DS-GVO verlangt jedoch, dass personenbezogene Daten nur so lange gespeichert werden, wie nötig.

Herausforderung: Die unbefristete Speicherung von Chatprotokollen oder Eingaben verstößt bei Verarbeitung personenbezogener Daten gegen die DS-GVO.

4) Datenübermittlung in Drittländer

Viele LLM-Anbieter haben ihre Server in den USA oder anderen Drittländern, bei denen das Datenschutzniveau nicht immer geklärt ist. Die DS-GVO erlaubt Übermittlungen in Drittländer nur unter strengen Bedingungen (z. B. US-EU Data Privacy Framework, Standardvertragsklauseln, Binding Corporate Rules). Stand heute ist der Datenschutz in den USA im freien Fall. Die Problemlage hier umfassend darzustellen, sprengt den Rahmen. Aktuell informierte Ende August 2025 die Federal Trade Commission (die Behörde, die in den USA für Recht und Ordnung im digitalen Raum zuständig ist) Apple, Google, Meta und andere darüber, dass sie die EU-Gesetze zum Digital Services Act (DSA) nicht unbedingt einhalten müssten, wenn dadurch die Ausdrucksfreiheit der US-Amerikaner:innen beeinträchtigt würde.

Herausforderung: Unzulässige Datenübermittlung in Drittländer kann zu Bußgeldern führen.

Analyse gem. RL zur Risikoanalyse und Risikobehandlung im Rahmen des KDM - Seite 3

Ollama ist grundsätzlich dazu geeignet diesen Herausforderungen zu begegnen, indem dieses Tool die Einbindung eines LLM unabhängig von den Servern großer Tech-Unternehmen ermöglicht.

Eine lokale Datenspeicherung ermöglicht es, Sprachmodelle wie Llama, Mistral oder Gemma direkt auf der eigenen Hardware oder in der internen IT-Infrastruktur zu betreiben. Dadurch bleiben alle Daten im Unternehmen, und werden nicht an externe Cloud-Anbieter übermittelt. Das ist ein zentraler Vorteil für die Einhaltung der DS-GVO, da personenbezogene Daten nicht an Dritte weitergegeben werden und die volle Kontrolle über die Datenverarbeitung besteht

Compliance-Vorteile: Durch den lokalen Betrieb entfällt die Abhängigkeit von externen Anbietern, was die Einhaltung der DS-GVO erleichtert. Unternehmen können so sicherstellen, dass keine Daten an Drittanbieter wie OpenAI oder andere Cloud-Dienste übermittelt werden. Ollama unterstützt die Installation einer DS-GVO-konforme KI-Lösung, die auf firmeneigenen Servern läuft, die idealerweise in Deutschland oder der EU stehen. Damit ist eine maximale Kontrolle über sensible Daten erreicht.

Zusammenfassung:

Ollama kann, sofern die Modelle lokal betrieben und korrekt konfiguriert werden, dazu beitragen den Datenschutz und die Datensouveränität beim Einsatz von LLM deutlich zu erhöhen.

Schlussbemerkung:

Angesichts dessen, dass mir bei der Informationsbeschaffung zu dieser Einschätzung nur sehr wenig allgemein verständliche Informationen zu Ollama zur Verfügung standen, habe ich einen befreundeten IT-Berater gebeten Ollama zu testen und die Eckdaten zusammen zu fassen. In der Anlage finden Sie seine Informationen.