

Nordrhein-Westfalen

KI-Sprachmodell für Hochschulen

[15.04.2024] Die Nutzung kommerzieller generativer KI-Lösungen wie ChatGPT & Co. ist datenschutzrechtlich schwierig. In Nordrhein-Westfalen startet ein vom Wissenschaftsministerium gefördertes Pilotprojekt, in dessen Rahmen ein KI-Sprachmodell speziell für Hochschulen entwickelt werden soll – auf Open-Source-Basis.

Bei der Nutzung Künstlicher Intelligenz spielt die Frage der Datensicherheit und der Datenhoheit für die Hochschulen Nordrhein-Westfalens eine entscheidende Rolle. Die Ruhr-Universität Bochum und die Universität zu Köln arbeiten jetzt an einem Pilotprojekt, um mittelfristig ein passgenaues Open-Source-Sprachmodell für eine große Hochschule – und perspektivisch mehreren Hochschulen – zu betreiben. Das berichtet das [Ministerium für Kultur und Wissenschaft des Landes Nordrhein-Westfalen](#), welches das Vorhaben mit rund 600.000 Euro fördert. „Die Anwendung Künstlicher Intelligenz muss ethischen Grundsätzen folgen. Dazu gehört auch der Schutz persönlicher Daten. Von der Förderung der Open-Source-KI profitieren alle Menschen, weil wir in dieser wichtigen Zukunftstechnologie Unabhängigkeit von kommerziellen Anbietern gewinnen“, sagte Wissenschaftsministerin Ina Brandes.

Open-Source-Sprachmodell bis 2025

Das wohl bekannteste und eines der leistungsstärksten KI-Sprachmodelle ist ChatGPT. Die Nutzung dieser kommerziellen Lösung geht für Hochschulen mit Herausforderungen einher: Neben der Frage der Kosten müssen der Datenschutz sowie Schulungsmöglichkeiten von Nutzerinnen und Nutzern geklärt werden. Eine Alternative zur Nutzung von ChatGPT und anderen kommerziellen Anbietern ist die konsequente Weiterentwicklung von Open-Source-Sprachmodellen durch die Hochschulen selbst.

Das Projekt Open-Source-KI.nrw soll dazu bis Ende 2025 erstmals den Aufbau einer vollständigen Infrastruktur eines Open-Source-Large-Language-Models (LLM) anhand eines konkreten Anwendungsfalls erproben. Im Fokus steht der Einsatz von Sprachmodellen in der Lehre. Studierende sollen in ausgewählten Kursen im Lern-Management-System über einen Chatbot mit dem Sprachmodell kommunizieren können. Der Chatbot greift für die Formulierung seiner Antworten sowohl auf seine allgemeinen Trainingsdaten zurück als auch auf konkrete Lerninhalte aus dem Kurs. So sollen die Antworten möglichst passend und hilfreich für die Lernenden sein.

Synergien mit Hochleistungs-Rechenzentrum

Die für das Projekt benötigten Rechenkapazitäten stellt das ebenfalls vom Ministerium geförderte Hochleistungsrechensystem an der Universität Köln zur Verfügung. So wird gleichzeitig mit dem Projekt erprobt, inwieweit die landeszentral verfügbare Hochleistungsrechner-Infrastruktur für solche Anwendungsfälle einsetzbar ist. Im Rahmen des Projekts soll außerdem in enger Kooperation mit dem ebenfalls vom Land geförderten Projekt [KI:edu.nrw](#) ein Kompetenz-Team zum Thema Open-Source-Sprachmodelle aufgebaut werden.

Die Projektergebnisse sollen allen Hochschulen zur Nachnutzung und Bewertung zur Verfügung stehen. Die Erfahrungen aus dem Betrieb der Software und der Infrastruktur werden über die etablierten Kompetenznetzwerke Moodle.NRW und HPC.NRW weitergegeben.

(sib)

Stichwörter: Hochschul-IT, KI, künstliche Intelligenz, Nordrhein-Westfalen