

Behörde / Institution
Landesbetrieb Information und Technik Nordrhein-Westfalen (IT.NRW)
Projektname
GovTeuken
Verwaltungsebene
Land
Website / URL
https://www.it.nrw/
Einreichungskategorie
Künstliche Intelligenz
Möchten Sie das Projekt noch in einer zweiten Kategorie einreichen?
Transformation
Projektbeschreibung
GovTeuken ist ein Open-Source-Projekt zur Entwicklung eines souveränen KI-Sprachmodells speziell für die öffentliche Verwaltung. Das von Juli 2025 bis Januar 2027 laufende Projekt wird durch den IT-Planungsrat gefördert und von IT.NRW gemeinsam mit dem Fraunhofer IAIS umgesetzt. Ziel ist ein nachnutzbares Domänen-Basismodell, das Verwaltungssprache, Dokumente, Aufgaben und Entscheidungskontexte besser versteht als allgemeine Sprachmodelle. Grundlage ist das in Deutschland entwickelte Soofi Basismodell 30B-A3B. Dieses wird durch Continued Pretraining (CPT) gezielt mit Verwaltungswissen erweitert und anschließend als Open-Source-Modell veröffentlicht. Für konkrete Anwendungsszenarien erfolgt zusätzlich eine Spezialisierung durch Supervised Fine-Tuning (SFT). Im Unterschied zu einzelnen KI-Pilotprojekten schafft GovTeuken eine wiederverwendbare Modellgrundlage für die Verwaltung. Durch das tief verankerte Verständnis von Verwaltungssprache können zukünftige Fachanwendungen schneller entwickelt, der Datenbedarf reduziert und Lösungen leichter auf andere Behörden übertragen werden. Für das Training wird eine umfangreiche Datenbasis aus öffentlichen Quellen, Verwaltungsdokumenten und bereitgestellten Datenspenden aufgebaut. Die Verarbeitung erfolgt unter Berücksichtigung von Datenschutz und Verwaltungsanforderungen. Für das Training kommen Hochleistungsrechner von WestAI/CLAIX mit bis zu 128 H100-GPUs sowie weitere GPU-Infrastrukturen zum Einsatz. Das Modell soll Verwaltungsmitarbeitende bei wissensintensiven Aufgaben unterstützen, beispielsweise durch Zusammenfassungen, Recherche, Textentwürfe, Klassifikation oder Frage-Antwort-Funktionen. Die fachliche Verantwortung und Entscheidung bleiben dabei weiterhin beim Menschen. Langfristig soll GovTeuken als souveräner KI-Baustein in föderale Infrastrukturen integriert und gemeinsam mit weiteren Verwaltungspartnern weiterentwickelt werden. Damit entsteht eine offene Grundlage für vertrauenswürdige und nachnutzbare KI in der öffentlichen Verwaltung.
Beschreiben Sie den technischen Innovationsgrad des Projektes und den Beitrag zur Verwaltungsmodernisierung.
GovTeuken ist technisch innovativ, weil es eine leistungsfähige und effizient betreibbare Open-Source-Modellgrundlage für die deutsche Verwaltung schafft. Die Architektur des Soofi 30B-A3B-Modells erlaubt eine effiziente Verarbeitung von Anfragen und ist im Schnitt 3,3x effizienter als vergleichbare Open-Source-Modelle. Für den Betrieb ist daher keine große Infrastruktur erforderlich. GovTeuken wird damit auch für kleinere Behörden, etwa auf kommunaler Ebene, relevant. Sie können das Modell perspektivisch selbst, in öffentlichen Infrastrukturen oder sogar verteilt auf Laptop-Endgeräten betreiben, ohne sich von externen Plattformen abhängig zu machen oder Datenschutzfragen an proprietäre Drittanbieter auszulagern. Der zweite Innovationskern liegt im Continued Pretraining. Dieses Verfahren wird in der Forschung genutzt, um Basismodelle für bestimmte Wissensgebiete zu verbessern (s. Shi et

al., <https://arxiv.org/abs/2404.16789>). GovTeuken überträgt diesen Ansatz auf die deutsche Verwaltungssprache mit Fachbegriffen, Rechtsbezügen, Zuständigkeiten, föderalen Ebenen, Verfahrenslogiken und Hierarchien. Das Modell soll diese Strukturen nicht nur über externe Dokumente nachschlagen, sondern im Training internalisieren. Dadurch entsteht eine Grundlage, die Verwaltungslogik, Textsorten und fachliche Zusammenhänge besser versteht als allgemeine Sprachmodelle.

Zur Verwaltungsmodernisierung trägt GovTeuken bei, indem es eine skalierbare Grundlage für KI-Anwendungen schafft. Das CPT-Modell und eine Blaupause zur Modellanpassung als Whitespacepaper, geplant für Q4/2026, befähigen Behörden und GovTech-Akteure, das Modell für eigene Aufgaben anzupassen, ohne auf große KI-Modelle aus den USA oder China aufzusetzen. Nachgelagerte Fachaufgaben müssen Verwaltungswissen damit nicht jeweils neu erlernen. Das reduziert Entwicklungsaufwand, Trainingskosten und Umsetzungsrisiken und verbindet moderne KI-Technologie mit effizientem Betrieb, Open Source und föderaler Nachnutzbarkeit.

Welchen ökonomischen Nutzen hat das Projekt?

GovTeuken erzeugt ökonomischen Nutzen vor allem durch Wiederverwendung, Skalierung und die Reduktion manueller Wissensarbeit. Viele Verwaltungsaufgaben bestehen aus Lesen, Prüfen, Zusammenfassen, Formulieren, Vergleichen und Recherchieren. Diese Tätigkeiten binden erhebliche Arbeitszeit, obwohl sie häufig wiederkehrenden Mustern folgen. Ein verwaltungsspezifisches Sprachmodell kann diese Arbeitsschritte vorbereiten und beschleunigen. Dadurch entstehen Zeitgewinne bei Sachbearbeitung, Fachreferaten, Serviceeinheiten und IT-nahen Rollen.

Zusätzlich reduziert GovTeuken langfristig Abhängigkeiten von proprietären KI-Angeboten und damit Lizenz-, Wechsel- und Integrationskosten. Eine offene Modellgrundlage kann mehrfach genutzt werden: in internen Assistenzsystemen, in Fachverfahren, in Wissensportalen, in Chat- und Recherchefunktionen sowie in Schulungs- und Unterstützungsprozessen.

Der ökonomische Mehrwert liegt auch in der besseren Nutzung vorhandenen Verwaltungswissens. Dokumente, Erlasse, Handreichungen, Prozessbeschreibungen und Fachinformationen werden durch in KI-Systeme (z. B. in NRW.Genius) integrierte Sprachmodelle schneller auffindbar und besser nutzbar. Das verringert Fehler, Doppelarbeit und Einarbeitungsaufwand. Gerade vor dem Hintergrund des demografischen Wandels kann GovTeuken helfen, Erfahrungswissen zu sichern und neue Beschäftigte schneller arbeitsfähig zu machen.

Schließlich stärkt das Projekt den Standort Deutschland und die Open-Source-Kompetenz. Öffentliche Mittel fließen nicht allein in geschlossene Einzellösungen, sondern in eine nachnutzbare, digitale Infrastruktur. Davon profitieren weitere Behörden, Forschung, Wirtschaft und die Community. GovTeuken schafft damit nicht nur Effizienz in einzelnen Prozessen, sondern einen Multiplikatoreffekt für souveräne KI-Kompetenz im öffentlichen Sektor.

Beschreiben Sie die Nachhaltigkeit der Lösung.

GovTeuken ist nachhaltig, weil es auf Wiederverwendbarkeit, offene Weiterentwicklung und langfristige Betriebsfähigkeit ausgelegt ist. Das Projekt adressiert keinen kurzfristigen Trend, sondern eine strukturelle Aufgabe der Verwaltung: den verantwortungsvollen Einsatz generativer KI unter Bedingungen von Datenschutz, Sicherheit, Nachvollziehbarkeit und digitaler Souveränität. Der Bedarf an verständlicher, schneller und qualitätsgesicherter Informationsverarbeitung wird in der Verwaltung dauerhaft bestehen.

Die Lösung ist so angelegt, dass sie nicht auf einen einzelnen Anwendungsfall beschränkt bleibt. GovTeuken kann als Modellgrundlage für unterschiedliche Assistenz- und Fachanwendungen dienen. Dadurch entsteht eine nachhaltige technische Basis, auf der neue Funktionen entwickelt, geprüft und betrieben werden können. Dokumentation, offene Schnittstellen und Open-Source-Komponenten erleichtern es anderen öffentlichen Organisationen, die Lösung zu verstehen, zu prüfen und an eigene Bedarfe anzupassen.

Nachhaltigkeit entsteht außerdem durch die Einbindung bestehender öffentlicher Strukturen. IT.NRW verfügt als zentraler IT-Dienstleister des Landes NRW über Erfahrung im sicheren Betrieb, in Integration, Skalierung und Betreuung verwaltungskritischer Systeme. Die

Zusammenarbeit mit MHKBD.NRW, Fraunhofer IAIS und dem SOOFI-Konsortium verbindet Verwaltungspraxis, Forschung und europäische KI-Entwicklung. So entsteht nicht nur Software, sondern auch Wissen, das in Betrieb, Support, Schulung und Weiterentwicklung überführt werden kann.

Das Projekt schafft Mechanismen, mit denen Verwaltungswissen gepflegt, aktualisiert und qualitätsgesichert in ein KI-Modell verankert werden kann. Damit bleibt die Lösung für spätere KI-Modelle anpassbar, wenn sich Rechtsgrundlagen, Verfahren oder organisatorische Anforderungen oder auch Modellarchitekturen ändern. Der offene Ansatz ermöglicht Rückflüsse in die Community, transparente Prüfung und kontinuierliche Verbesserung.

Wie trägt das Projekt zur Stärkung der Digitalen Souveränität bei?

GovTeuken stärkt digitale Souveränität, weil es der öffentlichen Verwaltung eine eigene, offene und kontrollierbare KI-Grundlage verschafft. Statt KI-Fähigkeiten von geschlossenen Plattformen oder einzelnen Anbietern abhängig zu machen, entsteht ein Open-Source-Modell, das geprüft, angepasst, betrieben und weiterentwickelt werden kann.

Verwaltungsrelevante Daten werden ausgewählt, gefiltert und, soweit erforderlich, anonymisiert. Datenspenden von IT.NRW werden für das Continued Pretraining aufbereitet; für konkrete Use Cases entstehen SFT-Daten als Question-Answer-Pairs. So wird die Entwicklung vom Datensatz bis zum spezialisierten Modell nachvollziehbar. Eine öffentliche Blaupause sichert die Übertragbarkeit.

GovTeuken ist darauf ausgerichtet, die Transparenzpflichten der europäischen KI-Verordnung zu erfüllen. Verwaltungen benötigen KI-Systeme, die leistungsfähig, erklärbar, prüfbar und rechtssicher betreibbar sind. Ein offenes, dokumentiertes und in Deutschland entwickeltes Modell ist tragfähiger als geschlossene Modelle aus den USA, China oder anderer Drittanbieter, deren Trainingsdaten, Änderungen und EU-Verfügbarkeit begrenzt kontrollierbar sind.

Dieses Risiko mindert GovTeuken durch Dokumentation und ganzheitliche Evaluation. Als möglichst objektive Bewertung wird das Modell deshalb u.a. auf dem MÖVE Datensatz der Bundesdruckerei ausgewertet werden (<https://moeve.bundesdruckerei.de/>). Europäische Werte wie Rechtsstaatlichkeit, Datenschutz, Nachvollziehbarkeit, Nichtdiskriminierung und demokratische Kontrolle werden in Datenaufbereitung, Training, Betrieb und Governance berücksichtigt und bewertbar.

Open Source erhöht die Prüfbarkeit durch Verwaltung, Wissenschaft und die Community. Mit dem IT-Planungsrat als Förderer und möglichem Partner und der Perspektive einer Einbindung in den D-Stack kann GovTeuken über NRW hinaus als föderal anschlussfähiger Baustein für souveräne KI in der Verwaltung wirken.

Projektwebsite

<https://www.mhkbd.nrw/presse-und-medien/pressemitteilungen/nordrhein-westfalen-spricht-moderne-verwaltung-sogar-mit-eigener-ki-sprache>

Umsetzungspartner

Leitung: Landesbetrieb Information und Technik Nordrhein-Westfalen (IT.NRW), Modellentwicklung: Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme IAIS, KI-Basismodell (lose assoziiert): SOOFI-Konsortium, Sovereign Open Source Foundational Models for European Intelligence

Zusätzliche Dokumente

Keine