

Behörde / Institution
Bundesdruckerei
Projektname
MÖVE
Verwaltungsebene
Bund
Website / URL
https://www.bundesdruckerei.de/de
Einreichungskategorie
Künstliche Intelligenz
Möchten Sie das Projekt noch in einer zweiten Kategorie einreichen?
Transformation
Projektbeschreibung
MÖVE ist der erste ganzheitliche KI-Modell-Vergleich, der speziell für die Anforderungen des öffentlichen Sektors entwickelt wurde. Es unterstützt Behörden und öffentliche Einrichtungen dabei, Entscheidungen für den verantwortungsvollen Einsatz von Künstlicher Intelligenz zu treffen. Jede Woche erscheinen neue Sprachmodelle (LLMs). Jedes beansprucht für sich, leistungsfähiger, sicherer oder effizienter zu sein als die Konkurrenz. KI-Tools bergen enormes Potenzial, um die öffentliche Verwaltung bürgernäher und zukunftsfähiger zu gestalten. Für Entscheidungstragende im öffentlichen Sektor entsteht daraus jedoch ein neues Problem: Die KI-Landschaft entwickelt sich schneller, als belastbare Bewertungsgrundlagen entstehen. Genau hier stoßen viele KI-Modell-Vergleiche an ihre Grenzen. Denn sie messen Fähigkeiten, wie englisches Textverständnis, mathematische Aufgaben oder allgemeines Weltwissen. Kaum berücksichtigt wird dagegen, was im Behördenalltag und für die deutsche Gesellschaft wichtig ist. Wer den Einsatz von KI in einer Behörde verantwortet, benötigt keine pauschale Modellempfehlung, sondern eine belastbare Entscheidungsgrundlage. MÖVE liefert dafür eine unabhängige und systematische Bewertung relevanter Sprachmodelle. Bewertet wird anhand von Kriterien, die für den öffentlichen Sektor und dessen Vertretern und Vertreterinnen des öffentlichen Sektors definiert werden. MÖVE schließt diese Lücke, indem es Sprachmodelle anhand verwaltungsnaher Aufgaben und relevanter Governance-Kriterien vergleichbar macht. Dazu gehören neben Leistungsfähigkeit u.a. Faktentreue (Halluzinationen), Transparenz, Nachhaltigkeit, Fairness und Sicherheit. Die Bewertung erfolgt über eine standardisierte und automatisierte Evaluierungspipeline, sodass neue Modelle, Kriterien und Datensätze kontinuierlich ergänzt werden können. Damit schafft MÖVE eine belastbare Grundlage für den verantwortungsvollen Einsatz von KI in Behörden. Das Projekt stärkt zugleich die digitale Souveränität der öffentlichen Verwaltung, weil Entscheidungen nicht allein auf Herstellerangaben oder internationalen Rankings beruhen, sondern auf nachvollziehbaren, unabhängigen und an europäischen Anforderungen ausgerichteten Bewertungen.
Beschreiben Sie den technischen Innovationsgrad des Projektes und den Beitrag zur Verwaltungsmodernisierung.
MÖVE weist einen hohen technischen Innovationsgrad auf, da im Rahmen des Projekts ein systematischer Bewertungsrahmen für Sprachmodelle im spezifischen Kontext der deutschen öffentlichen Verwaltung entwickelt wird. Anstatt allgemeine, internationale Benchmarks zu übernehmen, verbindet MÖVE technische Leistungsbewertung mit verwaltungsnahen Aufgaben, deutschen Sprach- und Fachkontexten sowie Governance-Kriterien wie Transparenz, Halluzinationen, Nachhaltigkeit, Fairness und Sicherheit. Die Evaluierung erfolgt über eine automatisierte Pipeline, die verschiedene Modelle anhand standardisierter Datensätze und Kriterien vergleichbar macht. Dadurch entsteht eine belastbare Grundlage, um Sprachmodelle nicht nur nach ihrer Leistungsfähigkeit, sondern auch nach ihrer Vertrauenswürdigkeit und Eignung für den öffentlichen Sektor zu bewerten. Das

Framework steht auf GitHub mit einer MIT-Lizenz allen zur Verfügung.

Der Beitrag zur Verwaltungsmodernisierung liegt insbesondere darin, dass MÖVE Orientierung für den verantwortungsvollen Einsatz von KI in Behörden schafft. Damit trägt MÖVE dazu bei, KI-Anwendungen effizienter und transparenter in Verwaltungsprozesse einzubetten.

Durch die einfache und offene Kommunikation der Projektergebnisse wird zudem der Kompetenzaufbau im KI-Bereich gefördert. Das Projekt stärkt zugleich die digitale Souveränität, da die Bewertungen nachvollziehbar, kontextspezifisch und an europäischen gesetzlichen Anforderungen und Werten ausgerichtet sind.

Welchen ökonomischen Nutzen hat das Projekt?

Der ökonomische Nutzen von MÖVE liegt darin, Investitionsentscheidungen beim Einsatz von Sprachmodellen in der öffentlichen Verwaltung fundierter und effizienter zu machen. Behörden stehen vor einer großen Auswahl an KI-Modellen, deren Leistungsfähigkeit, Kosten, Energieverbrauch und Eignung für konkrete Verwaltungsaufgaben stark variieren. MÖVE schafft hierfür eine vergleichbare Bewertungsgrundlage und hilft dabei, passende Modelle für spezifische Anwendungsfälle auszuwählen, anstatt pauschal auf große oder besonders bekannte Modelle zu setzen.

Langfristig trägt MÖVE dazu bei, Verwaltungsprozesse zu modernisieren, Personalressourcen gezielter einzusetzen und Skalierungspotenziale von KI-Anwendungen besser zu erschließen. Der wirtschaftliche Mehrwert entsteht somit nicht nur durch mögliche Effizienzgewinne, sondern auch durch mehr Transparenz, Risikoreduktion und eine bessere Steuerbarkeit öffentlicher KI-Investitionen.

Beschreiben Sie die Nachhaltigkeit der Lösung.

Die Nachhaltigkeit von MÖVE zeigt sich auf mehreren Ebenen: ökologisch, technisch und organisatorisch. Ökologisch trägt das Projekt dazu bei, den Ressourceneinsatz von KI-Systemen transparenter zu machen. MÖVE bewertet Sprachmodelle nicht nur nach ihrer Leistungsfähigkeit, sondern berücksichtigt auch Kriterien wie Energieverbrauch und CO₂-Äquivalente.

Dadurch wird sichtbar, welche Modelle für bestimmte Verwaltungsaufgaben ein gutes Verhältnis zwischen Qualität und Ressourcenverbrauch bieten. Dies unterstützt die Auswahl effizienterer Modelle und vermeidet den unnötigen Einsatz besonders großer und energieintensiver Systeme, wenn kleinere Modelle für den jeweiligen Zweck ausreichen.

Technisch ist MÖVE nachhaltig, weil die Lösung als wiederverwendbarer Bewertungsrahmen aufgebaut ist. Neue Modelle, Kriterien und Datensätze können kontinuierlich ergänzt werden, sodass die Lösung mit der dynamischen Entwicklung im KI-Bereich Schritt halten kann. Durch standardisierte Evaluierungsprozesse entsteht eine belastbare Grundlage, die langfristig für verschiedene Behörden und Anwendungsszenarien genutzt werden kann.

Organisatorisch stärkt MÖVE nachhaltige Verwaltungsmodernisierung, indem es Wissen, Transparenz und Vergleichbarkeit schafft. Behörden erhalten Orientierung für einen verantwortungsvollen KI-Einsatz und können Entscheidungen nachvollziehbar, wirtschaftlich und im Einklang mit europäischen Anforderungen treffen. So trägt MÖVE dazu bei, KI nicht kurzfristig experimentell, sondern dauerhaft verlässlich, ressourcenschonend und gemeinwohlorientiert in der Verwaltung einzusetzen.

Wie trägt das Projekt zur Stärkung der Digitalen Souveränität bei?

MÖVE stärkt die Digitale Souveränität, indem es der öffentlichen Verwaltung eine eigene, nachvollziehbare Bewertungsgrundlage für Sprachmodelle bereitstellt. Behörden sind dadurch weniger darauf angewiesen, sich ausschließlich auf Herstellerangaben oder internationale Rankings zu verlassen. Stattdessen können Modelle im spezifischen Kontext der deutschen Verwaltung geprüft werden.

MÖVE macht sichtbar, welche Modelle für konkrete Verwaltungsaufgaben geeignet sind und ob auch kleinere, offenere oder europäische Modelle eine leistungsfähige Alternative darstellen. Dadurch unterstützt das Projekt technologieoffene Entscheidungen, vermeidet Lock-in-Effekte und stärkt die strategische Handlungsfähigkeit der öffentlichen Hand.

Projektwebsite

https://moeve.bundesdruckerei.de/
Umsetzungspartner
BSI, Fraunhofer AISEC
Zusätzliche Dokumente
Keine