



Berlin, 01. April 2024

Gemeingut KI - Förderung von Open-Source basierten KI-Modellen (Bund)

Entwurf für einen Forderungskatalog

Wir werden auf Bundesebene **[Variante Berlin: in Berlin]*** die Entwicklung und den Einsatz von Open Source KI-Modellen fördern, um eine breite Nutzendenschaft von Schulen und Hochschulen, über öffentliche Verwaltung bis zu kleinen und mittelständischen Unternehmen in die Lage zu versetzen, KI-Modelle zu nutzen, sie an eigene Bedürfnisse anzupassen, sich an deren Weiterentwicklung zu beteiligen und sie auf Verzerrungen und Beschränkungen zu untersuchen. Unter Open Source KI-Modellen verstehen wir unter freien Lizenzen verfügbare, vertrauenswürdige und transparente KI-Systeme, die mit ebenfalls freien Trainingsdaten entwickelt werden, Anpassung, Weiterentwicklung und demokratische Kontrolle ermöglichen und gleichzeitig die KI-Kompetenz in verschiedenen Sektoren stärken. Der Fokus soll dabei vor allem auf großen Sprachmodellen mit allgemeinem Verwendungszweck liegen (sog. Large Language Models [LLM], General Purpose AI und Foundation Models). Die Bereitstellung und Kuration von freien Trainingsdaten wird gefördert.

Frei verfügbare, vertrauenswürdige und nachvollziehbare KI wird dabei dem Aufbau von KI-Kompetenz (ai literacy) einen Geschwindigkeitsschub geben. Es wird Transparenz, Nachvollziehbarkeit und Innovationskraft erhöht, aber auch die Grundlage für eine breite Verfügbarkeit und demokratische Kontrolle geschaffen.

Die volle Transparenz und Nachvollziehbarkeit von KI-Modellen soll die Grundlage für Standards für ethische Nutzung von KI sein. Durch die Offenlegung des Quellcodes und der verwendeten Trainingsdaten können Entwickler:innen, Forscher:innen und Nutzende die Funktionsweise und Entscheidungsfindung der Modelle besser verstehen. Dadurch können potenzielle ethische Bedenken frühzeitig erkannt und adressiert werden. Zudem ermöglicht Open Source eine breitere Beteiligung an der Entwicklung und Überprüfung von KI-Modellen, was zu einem vielfältigen Input und einem besseren Verständnis der Auswirkungen auf verschiedene Stakeholder führen kann.

Die Verfügbarkeit freier Sprachmodelle hat diverse Vorteile, wie die Unabhängigkeit von externen Dienstleistern, die Preise und Lizenzmodelle ändern oder Dienste einstellen können. Daten und Code können zudem innerhalb des eigenen Netzwerks verarbeitet werden, was die sichere Verarbeitung auch sensibler Informationen ermöglicht, zum Beispiel von Gesundheitsdaten in geschlossenen Räumen. Fachspezifische Open-Source-Modelle für verschiedene Anwendungsfälle wie für den deutschen Sprachraum oder die medizinische Forschung sind hoch spezialisiert.

Die Vision ist, eine aktive Community rund um Open Source Sprachmodelle (LLMs) zu schaffen, die neben einer Vielzahl von Open-Source Large Language Models auch eine Fülle hochwertiger Libraries und Tools bietet.

Die Bundesregierung **[Variante Berlin: Der Senat]** ergänzt konsequent die Förderprogramme für Innovation und Forschung auf Basis von Open-Source basierten KI-Modellen.

Es werden zudem Schulungen und Bildungsangebote für Schulen, Hochschulen und den Mittelstand gefördert, um die KI-Kompetenz (AI literacy) zu stärken. Open-Source-KI-Anwendungen sollen dabei in Lernplattformen integriert werden und unter freien Lizenzen ein Ökosystem für frei verfügbare Lern- und Trainingsmaterial im Sinne von Open Educational Resources geschaffen werden.

Der Bund **[Variante Berlin: Der Senat]** richtet eine zentrale Stelle als Kompetenzzentrum für KI-Anwendungen in der öffentlich Verwaltung ein, das Einsatz, Beschaffung und Weiterentwicklung von Open-Source KI-Technologien koordiniert, Leitlinien und Standards entwickelt und bei der Umsetzung von Pilot- und Leuchtturmprojekten unterstützt. Sie unterstützt auch bei der Bereitstellung benötigter Hardwareressourcen und Plattformen zur Nutzung und Weiterentwicklung der KI-Modelle und -Anwendungen. Dabei werden die Digitalisierungsagenturen wie ZenDiS, Sovereign Tech Funds und ITZ Bund eine zentrale Rolle einnehmen. **[Variante Berlin: ITDZ, Technologiestiftung Berlin und CityLAB]**

Die Datenlabore in den Bundesbehörden werden ertüchtigt, eigene Instanzen von Open-Source basierten KI-Modellen aufzusetzen, die Mitarbeitenden der Verwaltung zu schulen und Anreize für die Nutzung und Mitgestaltung zu setzen. Jedes Datenlabor erstellt einen Maßnahmenkatalog, wie KI-Technologien und -Kompetenzen auf Basis von Open-Source und freien Lizenzen in der Behörde aufgebaut werden. Auf europäischer und internationaler Ebene werden Kooperationsprojekte initiiert. **[Variante Berlin: (Absatz entfällt)]**

Begründung

Künstliche Intelligenz (KI) wird in rasanter Geschwindigkeit zur Basistechnologie für eine immer größere Zahl von Lebensbereichen. Gründe dafür sind unter anderem die möglichen Innovationen durch große Datenmengen und Rechenkraft, Effizienz- und Produktivitätssteigerungen durch Automatisierung und maschinelles Lernen, die Personalisierung von Diensten, sowie die Entwicklung neuer Technologien wie autonomem Fahren und neue medizinische Diagnostik.

Innovationen explodieren im Bereich der generativen KI, synthetische Biologie, Robotik, und mehr. Zeitgleich stehen große Herausforderungen vor uns, keine neuen Monopole entstehen zu lassen und die Technologie vertrauenswürdig, nachvollziehbar und sicher zu gestalten.

Wir werden sicherstellen, dass alle Menschen und gesellschaftlichen Akteure in Deutschland künftig ausreichenden Zugang und Kenntnisse im Bereich KI haben. Die Förderung von Open-Source-KI-Modellen zielt darauf ab, eine breite Nutzerschaft einzubeziehen, die KI-Modelle nutzen, anpassen, weiterentwickeln und ethische Fragen adressieren kann. Der

Schwerpunkt liegt auf großen Sprachmodellen mit universellem Nutzen, unterstützt durch die Bereitstellung und Kuratierung von freien Trainingsdaten in allen Sektoren.

Die volle Transparenz von Open-Source-KI-Modellen soll Standards für deren ethische Nutzung schaffen und die Entwicklung von KI-Kompetenz in allen Sektoren beschleunigen.

Die Verfügbarkeit freier Sprachmodelle bietet Vorteile wie Unabhängigkeit von externen Dienst Anbietern und ermöglicht die Verarbeitung sensibler Daten innerhalb des eigenen Netzwerks. Die Vision besteht darin, eine aktive Community rund um Open-Source-Sprachmodelle zu schaffen, die eine Vielzahl von Modellen sowie Libraries und Tools bietet.

* **Hinweis**: Antrag wird zweifach eingebracht, einmal in der Fassung für die Bundesebene, einmal in der Variante für das Land Berlin.

Aktuelle Informationen und Updates:

<https://spdnetz.de/pg-opensource/>

Kontakt:

Volkmar Stein

Co-Vorsitzender
SPD Forum Netzpolitik

v.stein@spdnetz.de

+49 173 855 2777

Twitter: @s_volkmar

Facebook: <http://fb.me/spdnetzpolitik>

www: <https://spdnetz.de>

SPD Landesverband Berlin

Kurt-Schumacher-Haus | Müllerstr. 163 | 13353 Berlin