Stand: 13.11.2026 11:54:58

Vorgangsmappe für die Drucksache 19/13785

"Sicherheit bei Künstlicher Intelligenz: Was ist zu tun für mehr Schutz, verantwortungsbewusste Forschung und Vorsorge?"

Vorgangsverlauf:

1. Initiativdrucksache 19/13785 vom 30.09.2026

Bayerischer Landtag

19. Wahlperiode 30.09.2026 **Drucksache** 19/**13785**

Antrag

der Abgeordneten **Florian von Brunn, Holger Grießhammer, Volkmar Halbleib, Anna Rasehorn, Doris Rauscher, Arif Taşdelen, Markus Rinderspacher, Nicole Bäumler, Christiane Feichtmeier, Ruth Müller, Horst Arnold, Martina Fehlner, Sabine Gross, Harry Scheuenstuhl, Dr. Simone Strohmayr, Ruth Waldmann, Katja Weitzel SPD**

Sicherheit bei Künstlicher Intelligenz: Was ist zu tun für mehr Schutz, verantwortungsbewusste Forschung und Vorsorge?

Der Landtag wolle beschließen:

Der Ausschuss für Wirtschaft, Landesentwicklung, Energie, Medien und Digitalisierung führt eine Anhörung gemäß § 173 Abs. 1 Satz 1 Geschäftsordnung für den Bayerischen Landtag (BayLTGeschO), hilfsweise gemäß § 173 Abs. 1 Satz 2 BayLTGeschO, zur Sicherheit Künstlicher Intelligenz (KI) und insbesondere sogenannter KI-Agenten (selbständige KI-Systeme) sowie zu notwendigen Schutz-, Forschungs- und Vorsorgemaßnahmen in Bayern, Deutschland und Europa durch.

Dabei sollen insbesondere folgende Fragen erörtert werden:

– Wie schnell entwickeln sich die Fähigkeiten fortgeschrittener KI-Systeme und KI-Agenten, und können Sicherheitsforschung, Schutzmaßnahmen und staatliche Kontrollmöglichkeiten mit dieser Entwicklung Schritt halten?

– Welche Lehren müssen aus den Sicherheitsvorfällen des Jahres 2026 gezogen werden? Welche Risiken entstehen, wenn KI-Agenten selbstständig kommunizieren und zusammenarbeiten, Schutzmechanismen umgehen oder auf fremde Systeme zugreifen?

– Welche Gefahren entstehen durch den Missbrauch von KI durch Kriminelle, Terrororganisationen und Staaten, insbesondere bei Cyberangriffen, Desinformation und Manipulation sowie militärischer Nutzung? Wie gut sind Kommunen und kritische Infrastruktur in Bayern auf KI-gestützte Cyberangriffe vorbereitet?

– Welche Gefahren entstehen durch KI für Datenschutz, Privatsphäre und Grundrechte, insbesondere durch staatliche oder private Überwachung, biometrische Erfassung, Persönlichkeitsprofile und „Social Scoring“, und reichen die bestehenden Schutzregeln und ihre Durchsetzung aus?

– Welche weitergehenden Risiken sind realistisch? Unter welchen Bedingungen könnten KI-Systeme eine Abschaltung verhindern, sich selbst Zugänge oder Ressourcen verschaffen oder Ziele gegen menschliche Eingriffe verfolgen?

– Welche Schutzmaßnahmen – wie getrennte (Test-)Umgebungen, klar begrenzte Zugriffsrechte, wirksame menschliche Kontrolle und Letztentscheidung, effektive Abschaltmöglichkeiten und klare Grenzen – sind für den Einsatz autonomer KI-Agenten, insbesondere in kritischer Infrastruktur, notwendig und wie sicher sind diese angesichts der rasanten Entwicklung?

– Reichen die KI-Verordnung der Europäischen Union und die bestehenden Sicherheits- und Aufsichtsstrukturen in Deutschland und Bayern aus? Wo bestehen mögliche Lücken, etwa bei Tests, unabhängigen Prüfungen, der Haftung der Anbieter,

Drucksachen, Plenarprotokolle sowie die Tagesordnungen der Vollversammlung und der Ausschüsse sind im Internet unter www.bayern.landtag.de - Dokumente abrufbar. Die aktuelle Sitzungsübersicht steht unter www.bayern.landtag.de – Aktuelles/Sitzungen zur Verfügung.

der Meldung und Untersuchung schwerer Vorfälle oder bei immer autonomeren KI-Agenten?

– Welche und wie viel Sicherheitsforschung ist notwendig und wie muss sie gefördert werden? In welchem Verhältnis sollten öffentliche Mittel für KI und für Forschung zu Sicherheit, Kontrolle und Zuverlässigkeit stehen? Sollte öffentliche KI-Förderung stärker als bisher mit Sicherheitsforschung verbunden werden, und welche unabhängigen Test- und Forschungsinfrastrukturen braucht Bayern?

– Wie kann verhindert werden, dass wirtschaftlicher Wettbewerbs- und Profitdruck zulasten notwendiger Sicherheitsmaßnahmen und Verbote geht? Braucht es unabhängige Institutionen, die KI-Systeme eigenständig prüfen sowie Sicherheitsrisiken unabhängig bewerten und veröffentlichen können?

– Was müssen Europa, Deutschland und Bayern konkret tun, um Sicherheit, technologische Leistungsfähigkeit und digitale Souveränität zu verbinden? Welche Sicherheitsstandards braucht insbesondere der Einsatz von KI-Agenten in Verwaltung, öffentlichen Unternehmen und kritischer Infrastruktur?

– Wie behält der Freistaat den Überblick über die in Landesbehörden, Kommunen und Hochschulen eingesetzten KI-Systeme, und wer beaufsichtigt ihren Einsatz mit der notwendigen Transparenz?

Begründung:

Künstliche Intelligenz entwickelt sich in hohem Tempo. Systeme können zunehmend selbstständig planen, digitale Werkzeuge einsetzen und autonom Aufgaben verfolgen. Damit stellt sich die Frage, ob Sicherheitsforschung und staatliche Kontrollmöglichkeiten mit dieser Entwicklung Schritt halten.

Die Ereignisse des Jahres 2026 zeigen, dass die Beantwortung dieser Frage immer dringlicher wird. Bei internen Sicherheitstests von OpenAI überwanden KI-Agenten Schutzmechanismen, verschafften sich Zugang zum Internet und griffen auf Systeme der Firma Hugging Face zu. Die Agenten fanden eigene Wege zur Kommunikation untereinander, tauschten Erkenntnisse aus, teilten Arbeit auf und entwickelten gemeinsam Strategien.

Bereits zuvor hatten OpenAI-Agenten das deutsche Programmierer-Wiki DseWiki mit über 15 000 Änderungen quasi zu einer eigenen Kommunikationsplattform gemacht. Sie tauschten dort Taktiken aus und koordinierten ihr Vorgehen. Auch bei Systemen anderer großer Anbieter – darunter Anthropic, Meta und Google – wurden während Sicherheitstests ähnliche, nicht vorgesehene Handlungen beobachtet.

Diese Vorfälle belegen, dass autonome Systeme bereits heute vorgesehene Grenzen überschreiten können. Zudem kann KI durch Kriminelle oder Staaten missbraucht werden. KI kann Aufgaben anders lösen als von Menschen beabsichtigt und mit wachsender Autonomie neue Kontrollprobleme schaffen. Hinzu kommen gesellschaftliche und staatliche Abhängigkeiten von wenigen hochleistungsfähigen Systemen, vor allem aus den USA. Außerdem erweitert KI die Möglichkeiten staatlicher und privater Überwachung, umfassender Persönlichkeitsprofile und der Bewertung von Menschen.

Ernstzunehmende Warnungen kommen aus der KI-Forschung selbst. Der Nobelpreisträger und KI-Pionier Geoffrey Hinton warnte im September 2026 mit Blick auf die verbleibende Zeit für wirksame Schutzmaßnahmen: „Vielleicht ein Jahr, aber nicht viel mehr als ein Jahr.“ Der Forscher Jacob Coxon, der drei Jahre bei OpenAI und Anthropic gearbeitet hatte, kündigte im September 2026 bei Anthropic und warf beiden Unternehmen vor, mit unser aller Leben zu spielen („gambling with our lives“). Selbst der Chefwissenschaftler von OpenAI, Jakub Pachocki, schrieb im gleichen Monat, kein KI-Konzern oder -Labor habe die Ausrichtung und Überwachung seiner Modelle bisher so unter Kontrolle, dass eine Weiterentwicklung im bisherigen Tempo verantwortbar sei. Er fordert verbindliche Sicherheitsstandards und internationale Koordinierung. Solche Warnungen zeigen, wie ernst wichtige Teile der Fachwelt mögliche Gefahren einschätzen.

Über Wahrscheinlichkeit und Ausmaß langfristiger Kontrollverluste oder sogar existenzieller Risiken besteht kein wissenschaftlicher Konsens. Gerade deshalb ist eine unabhängige Bewertung erforderlich: Welche Gefahren sind bereits nachweisbar, mit welchen Entwicklungen muss man rechnen und welche Vorsorge ist notwendig?

Mit der KI-Verordnung der Europäischen Union bestehen bereits Sicherheitsanforderungen für besonders leistungsfähige Modelle, die seit August 2026 von der Europäischen Kommission durchgesetzt werden können. Allerdings forderte die Staatsregierung 2025 ein Moratorium (Drs. 19/9388) – und Staatsminister für Digitales Dr. Fabian Mehring erklärte, Europa dürfe sich bei KI nicht selbst ausbremsen. Ministerpräsident Dr. Markus Söder verlangte auf der Social-Media-Plattform „X“ einen „Freedom Act für KI“. Vor dem Hintergrund der jüngsten Sicherheitsvorfälle und der öffentlichen Debatte ist zu klären, ob die bestehenden Sicherheitsanforderungen ausreichen, um schwerwiegende Risiken zu vermeiden.

Zugleich besteht ein gravierender Interessenkonflikt: Unternehmen, die von der schnellen Entwicklung und Einführung immer leistungsfähigerer Systeme wirtschaftlich profitieren, sind zugleich für deren Sicherheit verantwortlich. Deshalb ist zu klären, welche unabhängigen Prüfungs- und Aufsichtsstrukturen notwendig sind. Eine internationale Gruppe führender Forscher um Yoshua Bengio und Geoffrey Hinton hat bereits 2024 gefordert, mindestens ein Drittel der öffentlichen und privaten KI-Forschungsmittel für KI-Forschung und -Entwicklung in Sicherheit zu investieren.

Auch in Bayern besteht Handlungsbedarf. Eine teilweise Ausrichtung der mit der Hightech Agenda neu geschaffenen KI-Professuren auf Sicherheitsforschung ist nach Auskunft der Staatsregierung nicht erfolgt. Außerdem kann sie den Anteil der Sicherheitsforschung an den KI-Fördermitteln nicht angeben. Eine vollständige Erfassung der in Behörden, Kommunen und Hochschulen eingesetzten KI-Systeme lehnt sie in der Antwort auf eine Schriftliche Anfrage als unverhältnismäßigen Aufwand ab (siehe jeweils Drs. 19/9388). Sobald KI-Agenten in der Verwaltung selbstständig handeln, wird das jedoch zu einer wichtigen Transparenz- und Sicherheitsfrage.