

Verantwortungsvolle Nutzung Generativer KI, Rechtlicher Rahmen

Robert Haase

Diese Folien können unter den Bedingungen der [CC-BY 4.0](#) Lizenz wiederverwendet werden, falls nicht anders spezifiziert.

Hinweis

Ich bin weder Jurist, noch Datenschutzexperte. Ich kann nur Einblicke in rechtliche Rahmenbedingungen zur KI-Nutzung gewähren.

Im Zweifelsfall und bei konkreten Fragen und Projekten wenden Sie sich bitte an das Justitiariat der Universität.

justitiariat@zv.uni-leipzig.de

Justitiariat
Universität Leipzig
Ritterstraße 24
04109 Leipzig

KI-Kompetenztraining nach EU-AI Act

Die Anbieter und Betreiber von KI-Systemen ergreifen Maßnahmen, um nach besten Kräften sicherzustellen, dass ihr Personal und andere Personen, die in ihrem Auftrag mit dem Betrieb und der Nutzung von KI-Systemen befasst sind, über ein ausreichendes Maß an KI-Kompetenz verfügen, wobei ihre technischen Kenntnisse, ihre Erfahrung, ihre Ausbildung und Schulung und der Kontext, in dem die KI-Systeme eingesetzt werden sollen, sowie die Personen oder Personengruppen, bei denen die KI-Systeme eingesetzt werden sollen, zu berücksichtigen sind.
(aus EU AI Act, Art 4.)

KI-Kompetenztraining als Teil der Digitalisierungsstrategie

Zielvereinbarung

gemäß § 11 Absatz 2 SächsHSG

zwischen

der Universität Leipzig

vertreten durch die Rektorin Prof. Dr. Eva Inés Obergfell

und

dem Sächsischen Staatsministerium für Wissenschaft, Kultur
und Tourismus

vertreten durch den Staatsminister Sebastian Gemkow

für die Jahre 2025 bis 2028

1.1.6 Digitalisierung

Die UL setzt die formulierten strategischen Zielstellungen aus der Digitalisierungsstrategie des SMWK und der LRK für die Handlungsfelder IT-Infrastruktur und Dienste, administrative Hochschulprozesse um und entwickelt ein eigenes Umsetzungskonzept. In diesem verankert die UL operative Ziele, Meilensteine und Maßnahmen unter Berücksichtigung des gültigen Rechtsrahmens und der hochschulübergreifenden Zusammenarbeit und legt das Umsetzungskonzept bis zum 30.06.2026 dem SMWK vor.

Im Sinne von § 5 Absatz 2, Nummer 3 SächsHSG stärkt die UL (ohne Medizinische Fakultät) die digitalen und transformativen Kompetenzen¹ ihrer Beschäftigten in Verwaltung und Technik.

Dazu strebt sie für diese Beschäftigungsgruppe kumuliert für die Jahre 2025 bis 2028 eine Anzahl von 2.080 Teilnehmertagen an Fort- und Weiterbildungsveranstaltungen für diese Kompetenzen an.

KI-Anbieter und KI-Betreiber nach EU-AI Act

KI-Anbieter (Provider)

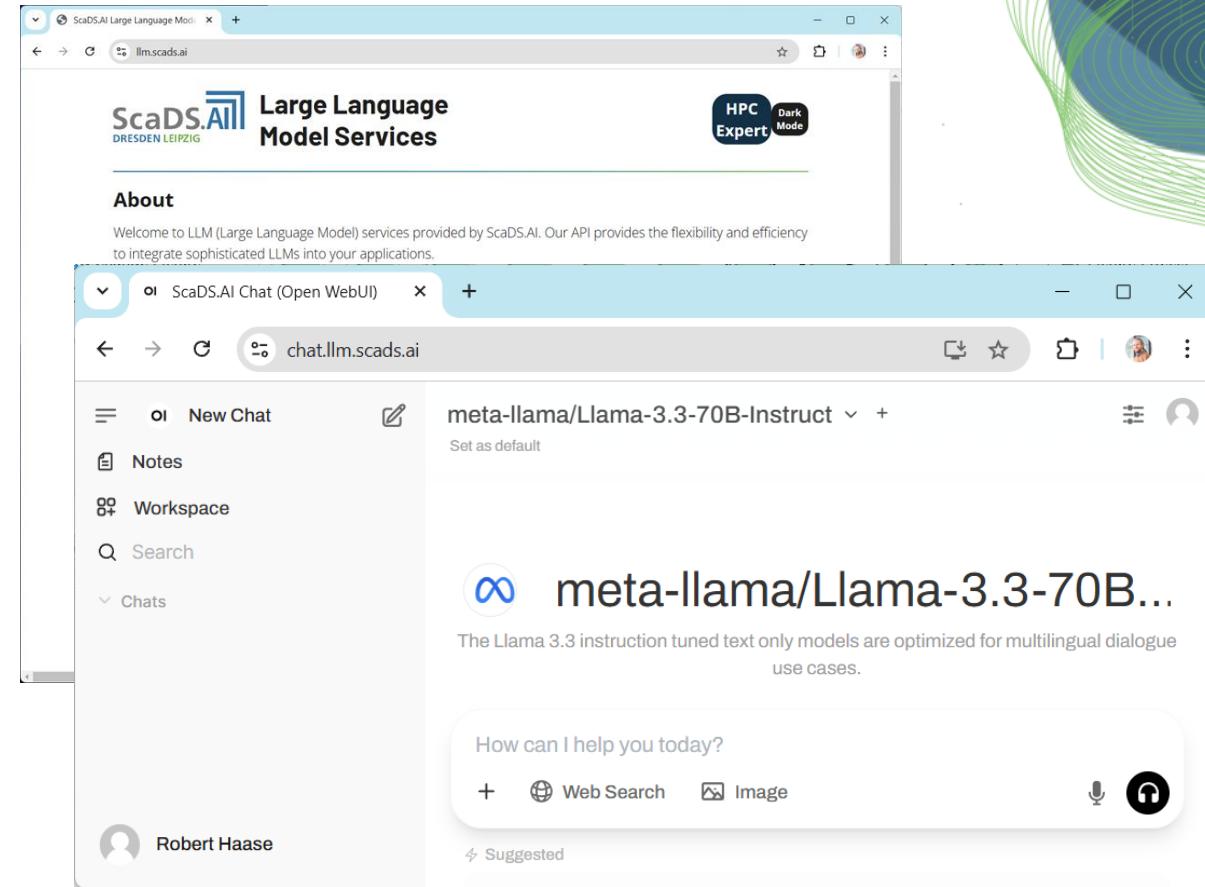
- Vermarktet einen Service oder ein Produkt auf Basis von KI-Systemen/-Modellen

KI-Betreiber (Deployer)

- stellt ein KI-System zu Eigennutzung unter eigene Aufsicht
- Ausnahme lt. EU AI Act: persönliche, nicht-professionelle Nutzung

Wissenschaftliche Forschung

- Weitgehend außerhalb der Betrachtung des EU AI Acts (Art 2)



<https://llm.scads.ai/> (nur aus VPN der TU Dresden)

URZ der UL baut
sowas gerade dank
ScaDS.AI Mitteln auf.



UNIVERSITÄT
LEIPZIG

KI-Anbieter und KI-Betreiber nach EU-AI Act

KI-Anbieter (Provider)

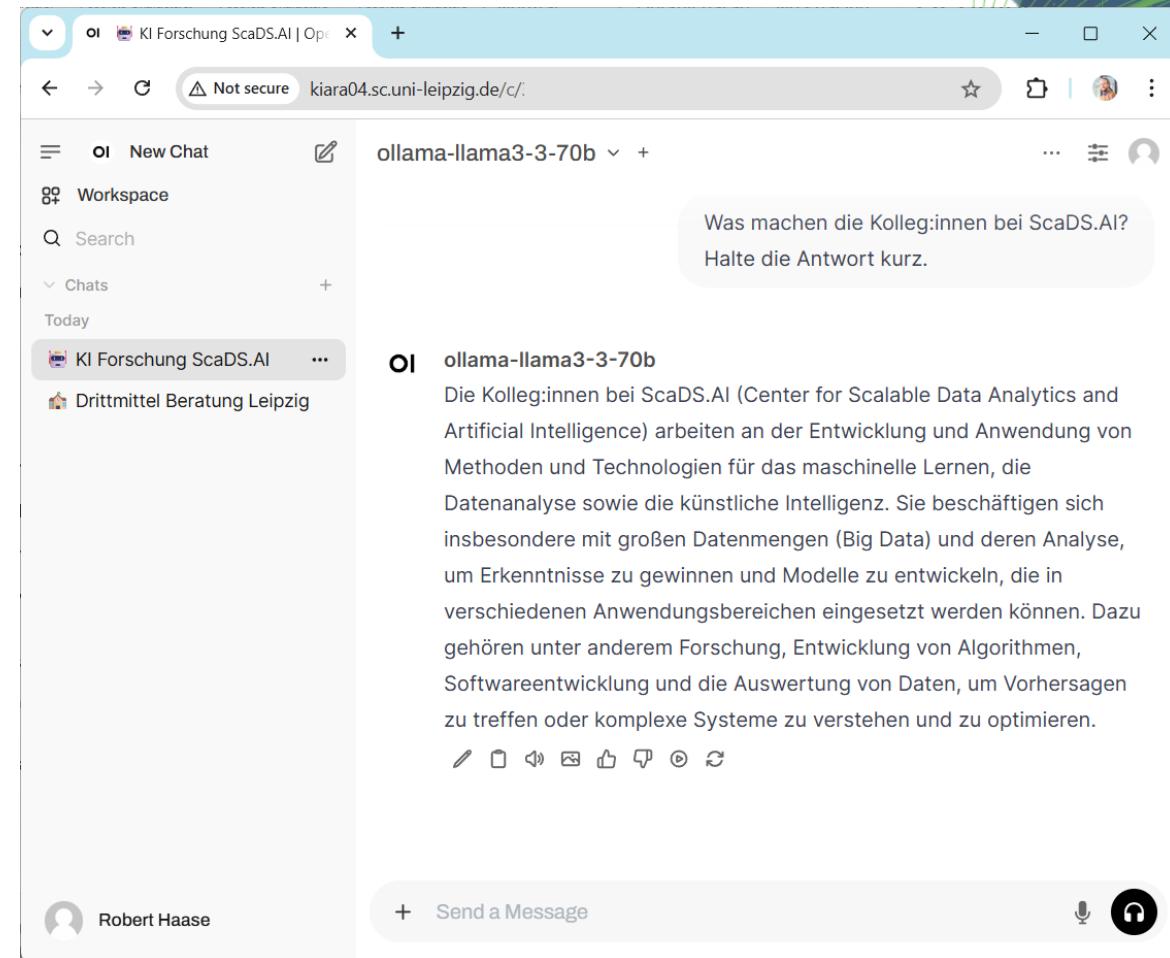
- Vermarktet einen Service oder ein Produkt auf Basis von KI-Systemen/-Modellen

KI-Betreiber (Deployer)

- stellt ein KI-System zu Eigennutzung unter eigene Aufsicht
- Ausnahme lt. EU AI Act: persönliche, nicht-professionelle Nutzung

Wissenschaftliche Forschung

- Weitgehend außerhalb der Betrachtung des EU AI Acts (Art 2)



URZ der UL baut
sowas gerade dank
ScaDS.AI Mitteln auf.



UNIVERSITÄT
LEIPZIG

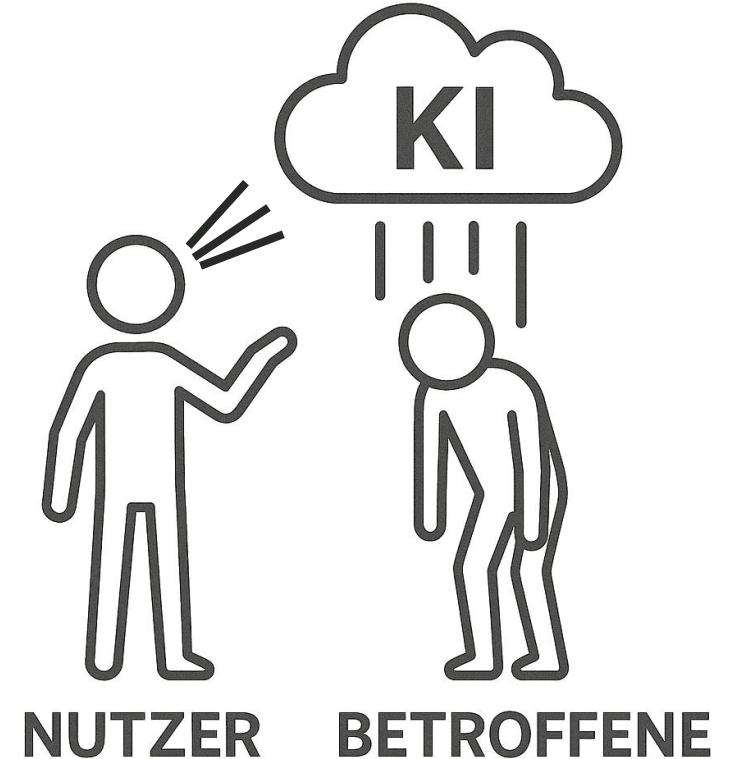
Nutzende und Betroffene

[Beruflich] Nutzende

- Menschen die KI einsetzen, um Daten zu verarbeiten
- Müssen im Umgang mit KI geschult sein

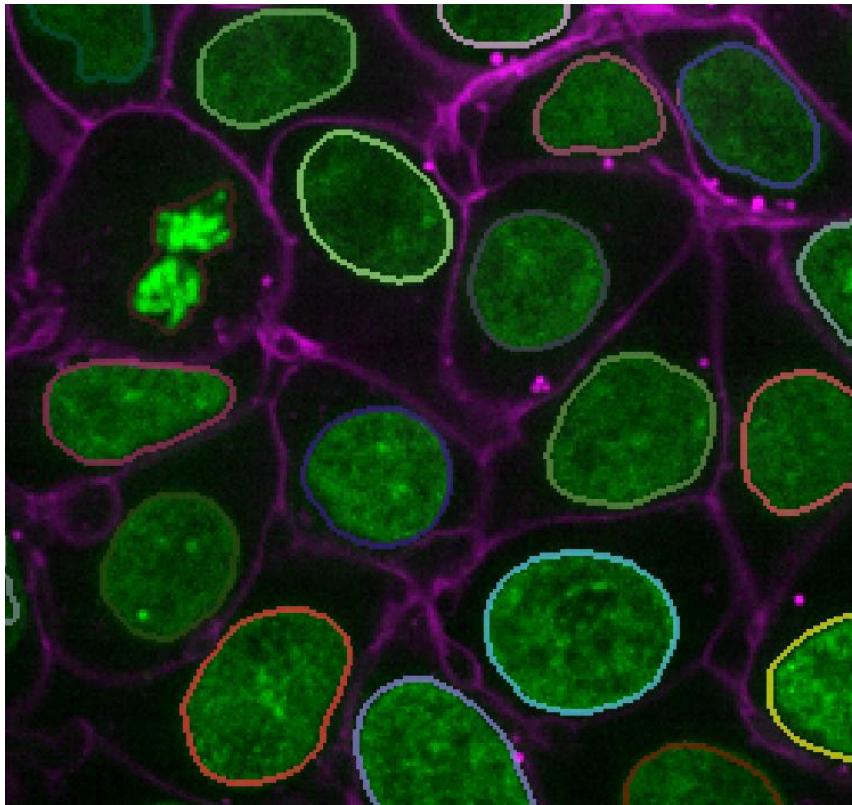
Betroffene

- Menschen deren Daten verarbeitet werden, oder die Konsequenzen der Verarbeitung spüren (direkt oder indirekt)



Dual use

Wofür können Algorithmen, bspw. für Bildsegmentierung in den Lebenswissenschaften, *noch benutzt werden?*



KI-Kompetenztraining nach EU-AI Act

56. „KI-Kompetenz“ die Fähigkeiten, die Kenntnisse und das Verständnis, die es Anbietern, Betreibern und Betroffenen unter Berücksichtigung ihrer jeweiligen Rechte und Pflichten im Rahmen dieser Verordnung ermöglichen, KI-Systeme sachkundig einzusetzen sowie sich der Chancen und Risiken von KI und möglicher Schäden, die sie verursachen kann, bewusst zu werden. (aus EU AI Act, Art 3)

Grundlagen /
Funktionsweise

Fachgerechte
Anwendung

Risiken / Rechtliche
Rahmenbedingungen

Teil 1

Teil 1/2

Teil 3

Richtlinie der Universität zum Einsatz von KI

Richtlinien der Deutschen Forschungsgemeinschaft zum Einsatz von KI (DFG)

- ▶ Es entspricht dem Berufsethos von Wissenschaftlerinnen und Wissenschaftlern, dass sie selbst für die Einhaltung der Grundprinzipien wissenschaftlicher Integrität einstehen. Der Einsatz generativer Modelle kann Wissenschaftlerinnen und Wissenschaftler von dieser inhaltlichen und formalen Verantwortung nicht entbinden.

Stellungnahme des Präsidiums
der Deutschen Forschungsgemeinschaft (DFG)
zum Einfluss generativer Modelle für die
Text- und Bilderstellung auf die Wissenschaften
und das Förderhandeln der DFG

September 2023

Richtlinien der Deutschen Forschungsgemeinschaft zum Einsatz von KI (DFG)

- Wissenschaftlerinnen und Wissenschaftler sollten bei der öffentlichen Zugänglichmachung ihrer Ergebnisse im Sinne wissenschaftlicher Integrität offenlegen, ob und welche generativen Modelle sie zu welchem Zweck und in welchem Umfang eingesetzt haben.

Stellungnahme des Präsidiums
der Deutschen Forschungsgemeinschaft (DFG)
zum Einfluss generativer Modelle für die
Text- und Bilderstellung auf die Wissenschaften
und das Förderhandeln der DFG

September 2023

Richtlinien der Deutschen Forschungsgemeinschaft zum Einsatz von KI (DFG)

- ▶ In wissenschaftlichen Publikationen können nur die verantwortlich handelnden natürlichen Personen als Autorinnen und Autoren in Erscheinung treten. Sie müssen sicherstellen, dass durch die Verwendung generativer Modelle kein fremdes geistiges Eigentum verletzt wird und kein wissenschaftliches Fehlverhalten etwa in Form von Plagiaten entsteht.
- ▶ Daraus folgt nach aktueller Einschätzung, dass der Einsatz von generativen Modellen bei der Antragstellung bei der DFG im Prozess der Begutachtung, Bewertung und Entscheidung als solcher grundsätzlich weder positiv noch negativ zu bewerten ist.

schaft (DFG)
für die
ssenschaften

Richtlinien der Deutschen Forschungsgemeinschaft zum Einsatz von KI (DFG)

- ▶ Bei der Erstellung von Gutachten ist der Einsatz von generativen Modellen mit Blick auf die Vertraulichkeit des Begutachtungsverfahrens unzulässig. Zur Begutachtung bereitgestellte Unterlagen sind vertraulich und dürfen insbesondere nicht als Eingabe für generative Modelle genutzt werden.

Stellungnahme des Präsidiums
der Deutschen Forschungsgemeinschaft (DFG)
zum Einfluss generativer Modelle für die
Text- und Bilderstellung auf die Wissenschaften
und das Förderhandeln der DFG

September 2023

Übung: Dürfen wir das?

Wir arbeiten als Reviewer für die DFG und würden gerne ein LLM einsetzen um

- Eine Zusammenfassung des Proposals zu generieren
- Rechtschreibfehler in unserem Review zu vermeiden

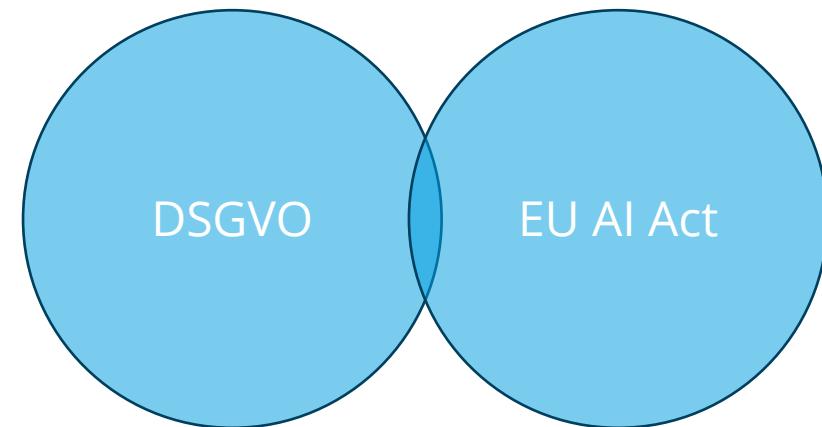
Datenschutz

(in aller Kürze)

Robert Haase

Datenschutz

- Recht auf Informationelle Selbstbestimmung, Teil des Persönlichkeitsrecht ([GG Art 2](#))
- Datenschutzgrundverordnung (DSGVO) -> einheitliche Regelungen EU-weit
- Gilt für Organisationen, die personenbezogene Daten von EU-Bürgern verarbeiten
- Ebenfalls relevant, aber nicht Datenschutz-Spezifisch:
EU AI Act



Grundprinzipien (Art 5 DSGVO)

- Rechtmäßigkeit, Verarbeitung nach Treu und Glauben, Transparenz
- Zweckbindung
- Datenminimierung
- Richtigkeit
- Speicherbegrenzung (zeitlich)
- Integrität und Vertraulichkeit (Datensicherheit)
- Rechenschaftspflicht der datenverarbeitenden Stellen

Wir dürfen Personen-bezogene Daten nur mit KI-System verarbeiten, wenn Zustimmung vorliegt und die Daten erforderlich sind

Daten nicht erfassen oder löschen, wenn Grund zur Speicherung nicht (mehr) gegeben.

-> Man darf keine Daten sammeln für den Fall, dass KI später was damit anfangen könnte!

Rechte der betroffenen Personen

- Auskunftsrecht
 - Bspw. [KI-gestützte] Entscheidungsprozesse müssen erklärbar / nachvollziehbar sein. (Siehe auch EU-AI Act)
- Recht auf Berichtigung und Löschung („Recht auf Vergessenwerden“)
- Recht auf Einschränkung der Verarbeitung
- Recht auf Datenübertragbarkeit
- Widerspruchsrecht
- Recht auf Beschwerde bei Datenschutzbehörden

Explainable AI ist ein Forschungsgebiet. Viele Entscheidungen von KI können wir nicht erklären

Rechtmäßigkeit der Verarbeitung ([Art. 6 DSGVO](#))

Verarbeitung erlaubt...

- mit Einwilligung der Betroffenen
- zur Vertragserfüllung auf Anfrage der Betroffenen
- wegen rechtlichen Verpflichtungen der Daten-Verarbeitenden
- wenn lebenswichtig für Personen
- wenn im öffentlichen Interesse / Ausübung öffentlicher Gewalt
- berechtigtes Interesse (v. Dritten) wenn Grundrechte/-freiheiten nicht verletzt (Besonderheiten bzgl. Behörden und/oder Kinder).

Pflichten für Unternehmen

gelten auch für Universitäten
im entsprechenden Sinne

- Bestellung eines Datenschutzbeauftragten (falls erforderlich)
- Datenschutz-Folgenabschätzung (DSFA) bei risikoreicher Verarbeitung
- Abschluss von Auftragsverarbeitungsverträgen (AVV)
- Meldung von Datenschutzverletzungen innerhalb von 72 Stunden
- Dokumentation und Nachweis der Datenschutzkonformität (Verzeichnis der Verarbeitungstätigkeiten)
- **Schulung und Sensibilisierung von Mitarbeitenden**

Beispiel: Microsoft Copilot

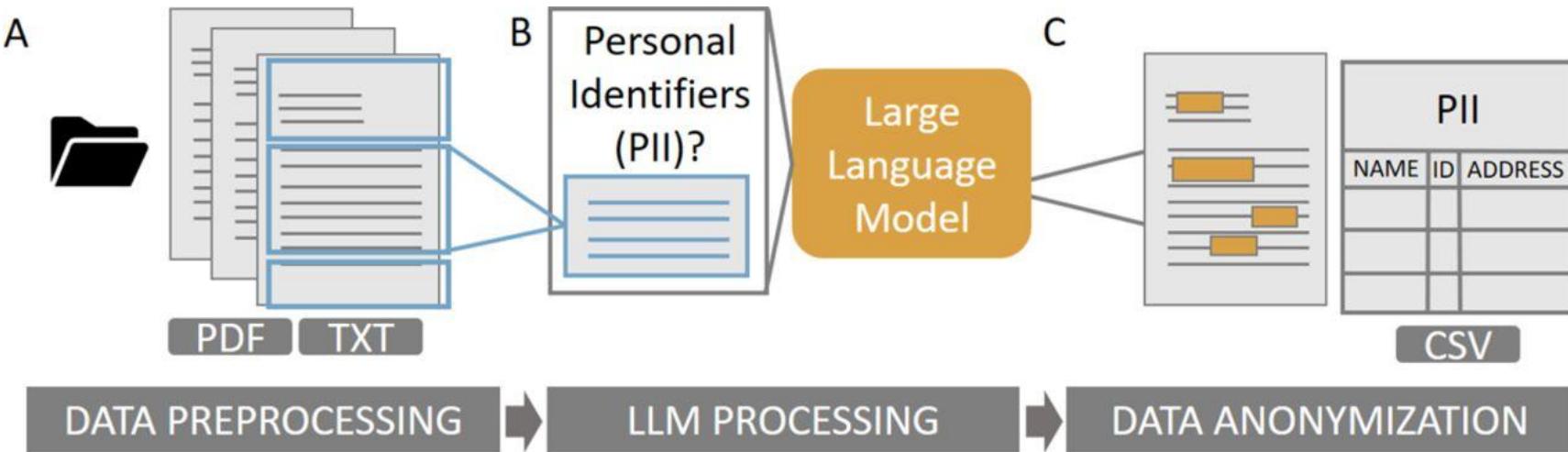
Generative KI in Office-Produkten

The screenshot shows a Microsoft Learn page titled 'Daten, Datenschutz und Sicherheit für Microsoft 365 Copilot'. The page includes a sidebar with navigation links like 'Nach Titel filtern', 'Bereitstellen und Aktivieren von Microsoft 365 Copilot', and 'Maximieren des Werts von Microsoft 365 Copilot'. The main content area features a large heading and several sections: 'In diesem Artikel', 'Wie verwendet Microsoft 365 Copilot Ihre geschützten Organisationsdaten?', and 'Zusätzliche Ressourcen' which lists 'Ereignisse', 'Ki-Fähigkeiten fest Herausforderung', 'Training', and 'Zertifizierung'.

The screenshot shows a Microsoft Learn page titled 'Wie verwendet Microsoft 365 Copilot Ihre geschützten Organisationsdaten?'. It features a sidebar with links like 'Nach Titel filtern', 'Bereitstellen und Aktivieren von Microsoft 365 Copilot', and 'Maximieren des Werts von Microsoft 365 Copilot'. The main content area contains a large heading, a 'Wichtig' section with a note about data usage for training, and a 'Transparenzhinweis für Microsoft 365 Copilot' section at the bottom.

Anonymisierung / Pseudonymisierung

- Wichtig für Personenbezogene Daten, bspw. wenn Datenanalyse auch ohne Klarnamen möglich ist.
- Schlüssel für Prozessierung human-medizinischer Daten

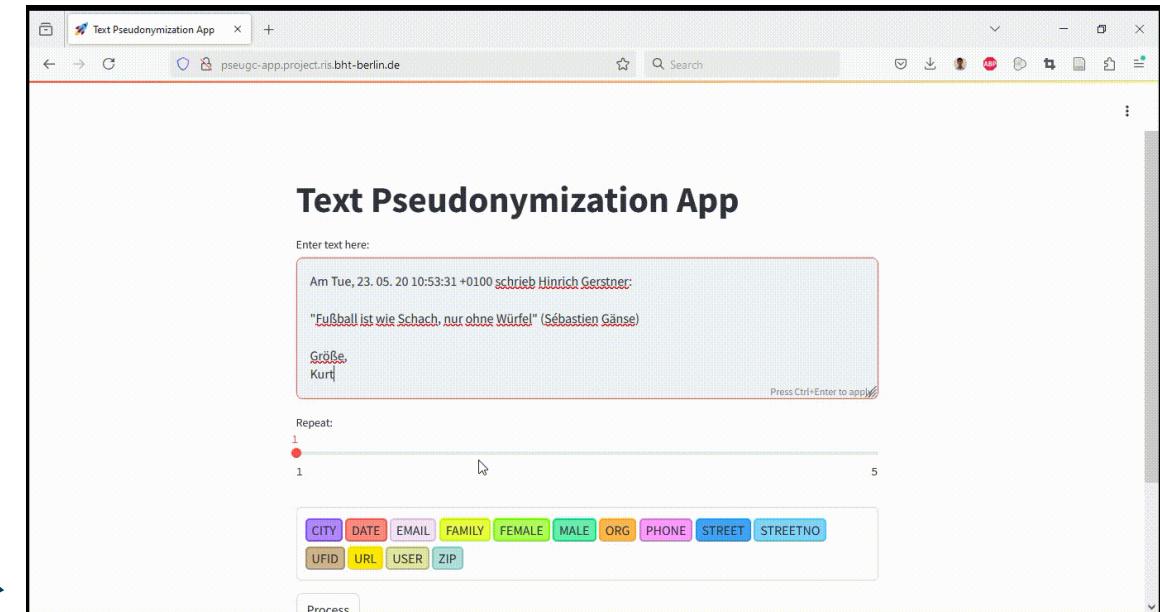


“success rate of 98.05% in removing text characters carrying personal identifying information” Wiest et al. (2024)

Anonymisierung / Pseudonymisierung

- Pseudonymisierung erforderlich für Einsatz in der Praxis
- For improving models, training data is not available at the moment
- "we argue that [...] better de-identification and pseudonymization tools are a prerequisite for responsible usage and research of [...] LLMs, in health care" (Saha & Biesmann 2025)

Open Source



KI-Verordnung der Europäischen Union “EU AI Act”

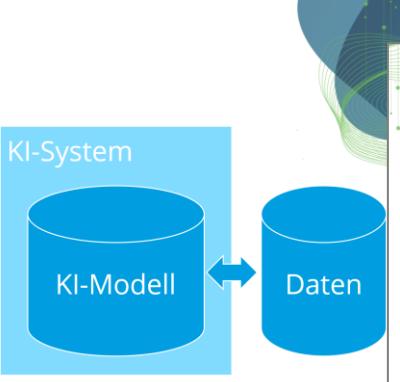
Robert Haase

EU AI Act - Definitionen

KI-Modelle
KI-Systeme
KI-Betreiber
KI-Anbieter
Nutzende
Betroffene

Künstliche Intelligenz (KI)

- KI-Modelle** (bspw. Sprachmodelle):
 - Text-to-text
 - Bildgenerierung, -interpretation,...
 - Beispiele: GPT-4o, Gemini, Llama, DeepSeek, Teukon...
- KI-Systeme**: Kombination von [Sprach]modellen mit:
 - Web-Suche,
 - Dokumentenmanagement,
 - Datenbanken, ...
 - Beispiele: ChatGPT, Perplexity, You.com



ScaDS.AI
DRESDEN LEIPZIG
KI-Kompetenzen
Robert Haase
@haesleinhuepf
26. Mai 2025

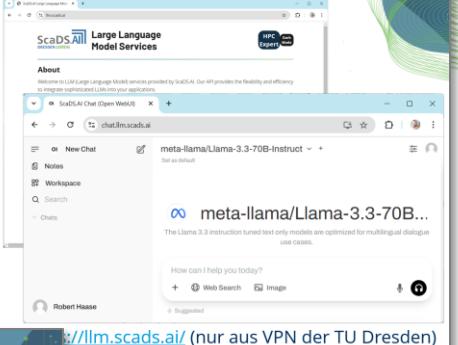
Siehe auch Art 3 EU AI Act: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32024R1689>

KI-Anbieter und KI-Betreiber nach EU-AI Act

KI-Anbieter (Provider)
• Vermarktet einen Service oder ein Produkt auf Basis von KI-Systemen/-Modellen

KI-Betreiber (Deployer)
• stellt ein KI-System zu Eigennutzung unter eigene Aufsicht
• Ausnahme lt. EU AI Act: persönliche, nicht-professionelle Nutzung

Wissenschaftliche Forschung
• Weiterabstand zu Berührungszone



6 TECHNISCHE UNIVERSITÄT DRESDEN
7 UNIVERSITÄT LEIPZIG

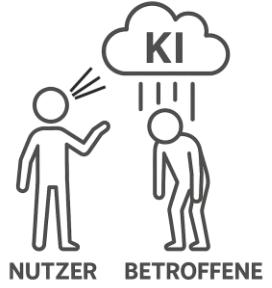
Nutzende und Betroffene

[Beruflich] Nutzende

- Menschen die KI einsetzen, um Daten zu verarbeiten
- Müssen im Umgang mit KI geschult sein

Betroffene

- Menschen deren Daten verarbeitet werden, oder die Konsequenzen der Verarbeitung spüren (direkt oder indirekt)



ScaDS.AI
DRESDEN LEIPZIG
KI-Kompetenzen
Robert Haase
@haesleinhuepf
26. Mai 2025

Abbildung wurde erzeugt mit ChatGPT (und nachbearbeitet)

Geltungsbereich EU AI Act

Laut Article 2 (Scope):

- Anbieter, Betreiber, Importierende, Distributoren von KI-Systemen/Modellen innerhalb und außerhalb der EU, solange EU in naheliegender Form betroffen sind, bspw. KI-System produziert Text innerhalb der EU.
- Produktanbieter, die KI in Produkten nutzen.
- Betroffene Personen innerhalb der EU.

Geltungsbereich EU AI Act

Auszug EU AI Act Article 2 (Scope):

“6. This Regulation does not apply to AI systems or AI models, including their output, specifically developed and put into service for the sole purpose of scientific research and development.”

“10. This Regulation does not apply to obligations of deployers who are natural persons using AI systems in the course of a purely personal non-professional activity.”

“12. This Regulation does not apply to AI systems released under free and open-source licences, unless they are placed on the market or put into service as high-risk AI systems or as an AI system that falls under Article 5 or 50.”

5: Verbotene KI-Praktiken

50: Transparenzpflichten

Quiz: Einordnung der Universität

Was sind “wir” im Kontext des EU AI Acts?

Anbieter



Betreiber



Nutzende

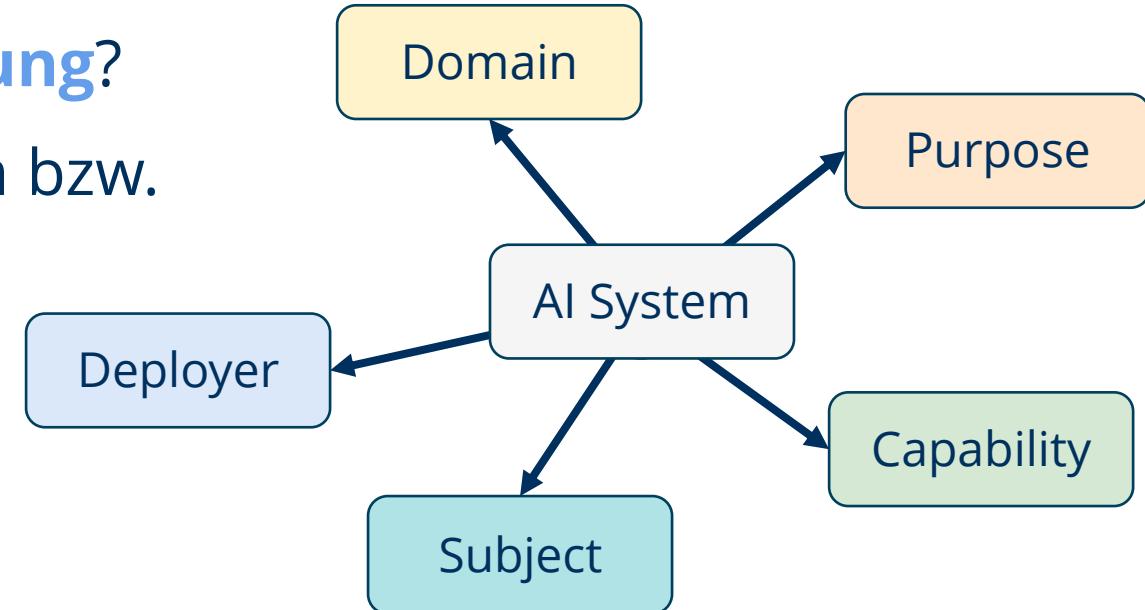


Betroffene



KI-bezogene Risiken

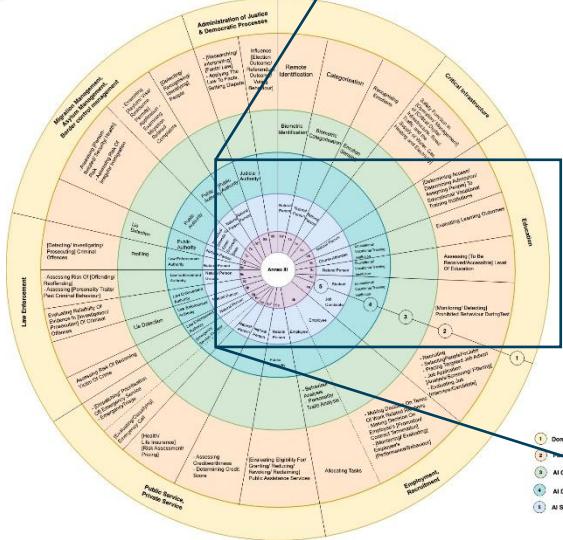
- In welcher **Domäne** wird das KI-System benutzt?
- Was ist der **Zweck** des KI-Systems?
- Was sind die **Fähigkeiten** des KI-Systems?
- Wer **stellt** das KI-System **zur Verfügung**?
- Wer ist **betroffen** von dem KI-System bzw. dessen Einsatz?



KI-bezogene Risiken

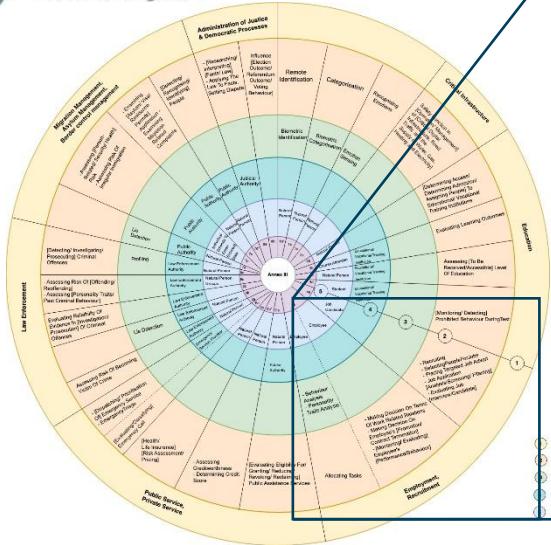
Zulassung

- 1 Domain**
 - 2 Purpose**
 - 3 AI Capability**
 - 4 AI Deployer**
 - 5 AI Subject**



KI-bezogene Risiken

- 1 Domain
 - 2 Purpose
 - 3 AI Capability
 - 4 AI Deployer
 - 5 AI Subject



Task-Zuweisungen

Beförderungen, Entlassungen

Einstellungen (Bewerberauswahl)

- Making Decision On Terms Of Work Related Relations
- Making Decision On Employee's [Promotion/ Contract Termination]
- [Monitoring/ Evaluating] Employee's [Performance/Behaviour]

Allocating Tasks

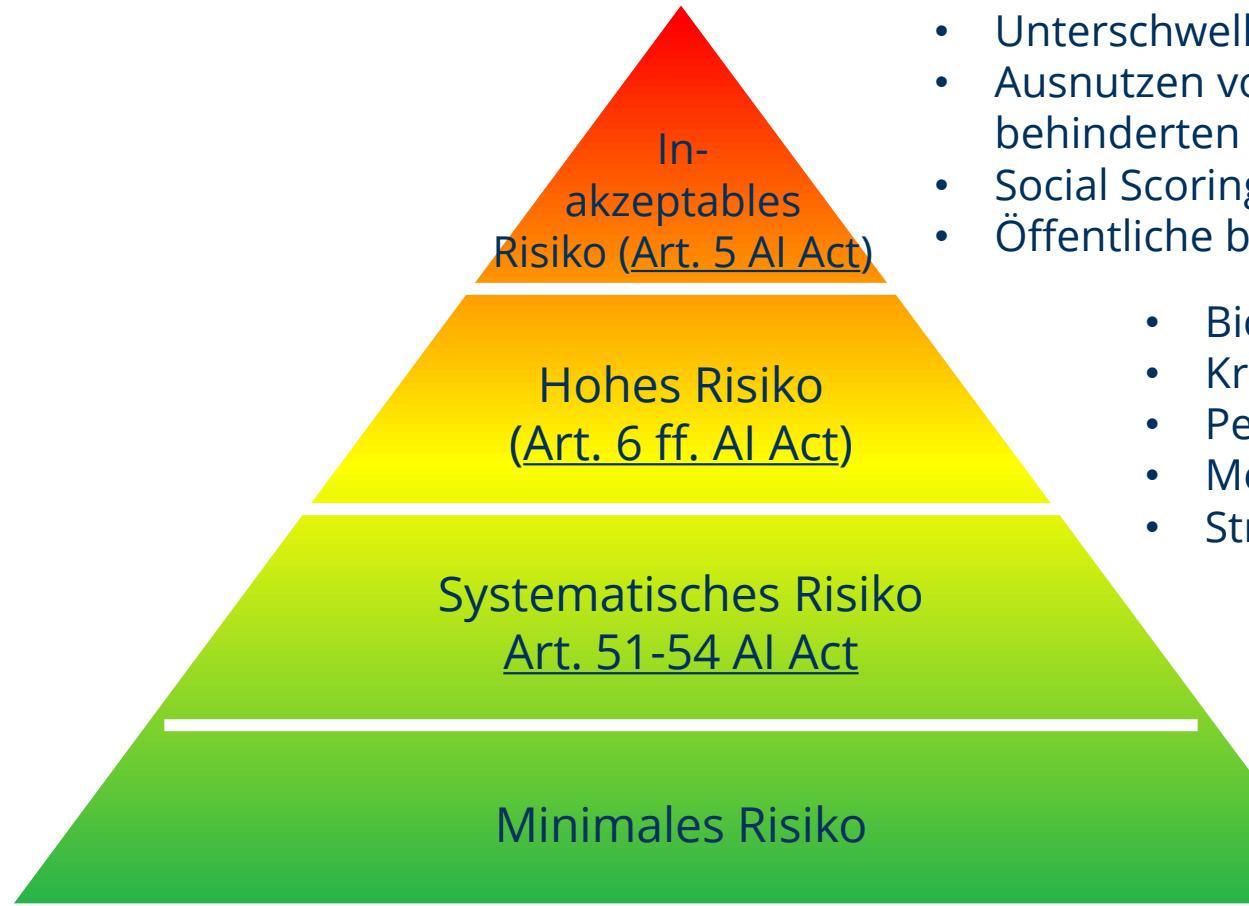
Employment,
Recruitment

Modifiziert von AI Risk Ontology (draft), licensed
[CC-BY 4.0](https://delaramglp.github.io/airo/) <https://delaramglp.github.io/airo/>

EU AI Act – Risikobasierter Ansatz

Kategorisierung von KI-Systemen nach Risiko

Konsequenzen



- Unterschwellige Beeinflussung
- Ausnutzen von Kindern oder gestig behinderten Menschen
- Social Scoring
- Öffentliche biometrische Identifikation

- Biometrie
- Kritische Infrastruktur
- Personalwesen
- Medizinprodukte
- Strafverfolgung

- Chatbots
- Deepfakes

- Computerspiele
- Spam-Filter
- Produktvorschläge

Qualitätssicherung / Externes Review

Risiko-management

Transparenz-pflicht

Menschliche Aufsicht

Dokumentations-pflicht

Ggf. Datenschutz-richtlinien beachten



Dokumentationspflicht / Transparenz

Nutzer müssen

- KI-gestützte Entscheidungsprozesse nachvollziehen können (Rückverfolgbarkeit),
- Entscheidungen anfechten können,
- der Verarbeitung ihrer Daten widersprechen können.



Anbieter / Hersteller müssen

- Algorithmen dokumentieren,
- inkl. Risikobewertung,
- Entscheidungen durch Menschen überwachen.

Sprachmodelle machen Fehler

(nach wie vor)



Translate this to german: A general-purpose AI model shall be presumed to have high impact capabilities pursuant to paragraph 1, point (a), when the cumulative amount of computation used for its training measured in floating point operations is greater than 10^{25} .

Ein Allzweck-KI-Modell gilt gemäß Absatz 1, Punkt (a), als mit hohen Auswirkungen versehen, wenn die kumulative Menge der für sein Training verwendeten Berechnungen, gemessen in Gleitkommaberechnungen (FLOPs), größer als 10^{15} ist.

Der Nutzer des Sprachmodells
/ Autor des übersetzten Textes
ist verantwortlich

Sprachmodelle können Fehler korrigieren

Bist Du Dir sicher dass die Übersetzung korrekt ist?

Diese Frage kann man eigentlich immer stellen.

Gute Frage – lass uns das nochmal ganz genau anschauen. Hier ist der Originalsatz:

A general-purpose AI model shall be presumed to have high impact capabilities pursuant to paragraph 1, point (a), when the cumulative amount of computation used for its training measured in floating point operations is greater than 10^{25} .

Und hier eine präzisere, wörtlichere und korrektere Übersetzung:

Ein Allzweck-KI-Modell gilt gemäß Absatz 1, Buchstabe a, als mit hohem Wirkungspotenzial versehen, wenn die kumulative Rechenleistung, die für sein Training verwendet wurde – gemessen in Gleitkommaoperationen (Floating Point Operations) – mehr als 10^{25} beträgt.

Sprachmodelle können Fehler korrigieren

🔍 Erläuterung:

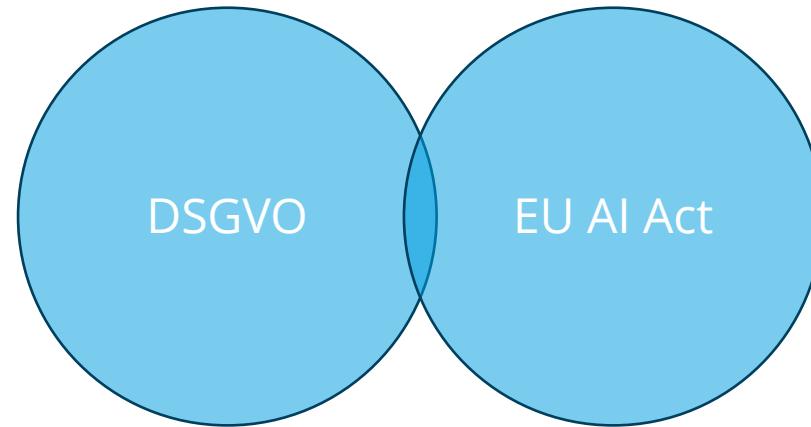
- "General-purpose AI model" → "Allzweck-KI-Modell"
- "shall be presumed to have high impact capabilities" → "gilt als mit hohem Wirkungspotenzial versehen"
- "pursuant to paragraph 1, point (a)" → "gemäß Absatz 1, Buchstabe a"
- "cumulative amount of computation used for its training measured in floating point operations" → "kumulative Rechenleistung ... gemessen in Gleitkommaoperationen"
- "greater than 10^{25} " → "mehr als 10 hoch 25"

Wenn du magst, kann ich dir auch eine etwas freiere, flüssigere Version machen.



Fazit

- EU AI Act und DSGVO ergänzen sich gegenseitig im Schutz von Grundrechten
- Schlüsselemente: Risiken, Verantwortlichkeit, Transparenz und menschliche Aufsicht



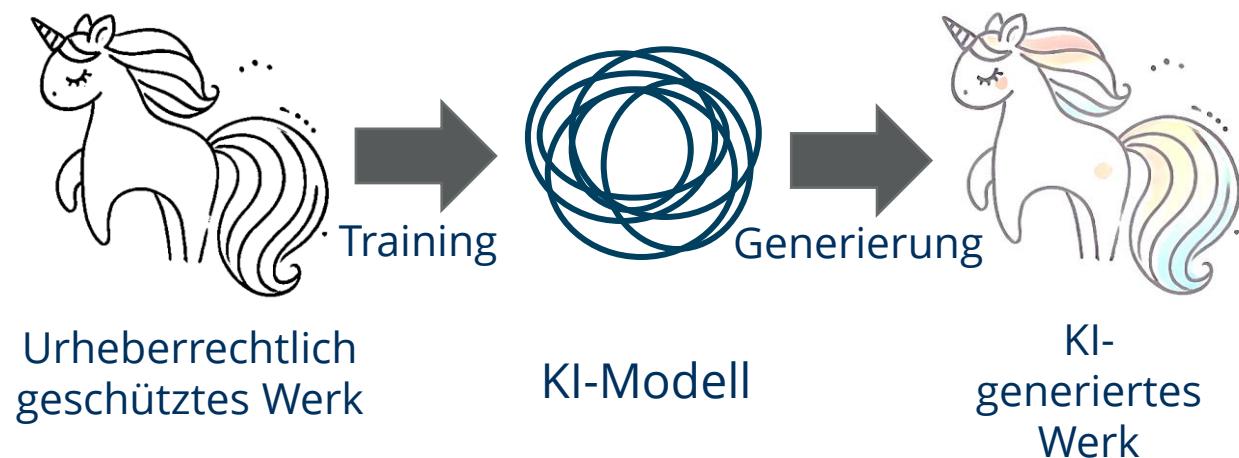
Urheberrecht im Kontext Generativer KI

Robert Haase



Urheberrecht

- Problem: LLMs sind trainiert mit geschützten Werken und geben diese mitunter wieder
- LLM-Training kann als Form von *Text und Data Mining* rechtlich legitim sein ([§44 UrhG](#))
- Autor von KI-generierten Werken verantwortlich ggf. Rechte einzuholen (schwierig)
- Schöpfungshöhe beachten



Urheberrecht

- Vermutlich haben KI-Anbieter bestimmte Nutzungsrechte erworben und andere nicht.
- KI-Anbieter verletzen die Rechte anderer aber auch immer wieder.
- Siehe auch:
Markenschutzrecht



KI-generierte Plagiate

Beim Generieren von Text ist man
davor nicht sicher...
... abhängig vom Anwendungsfall

- ChatGPT 4o 

Formuliere einen kurzen Text zum EU AI Act.



Der EU AI Act ist das erste umfassende Gesetz zur Regulierung von Künstlicher Intelligenz in der Europäischen Union. Ziel ist es, Innovation zu fördern und gleichzeitig Risiken für die Sicherheit, Grundrechte und demokratischen Werte zu minimieren. Das Gesetz verfolgt einen risikobasierten Ansatz: Je höher das potenzielle Risiko einer KI-Anwendung, desto strenger die Anforderungen. Besonders reguliert werden etwa biometrische Überwachung, KI in der Strafverfolgung oder Systeme mit manipulativem Einfluss. Der EU AI Act soll ein einheitliches Regelwerk für alle Mitgliedstaaten schaffen und als globales Vorbild für verantwortungsvollen KI-Einsatz dienen.

Wie könnte
denn hier
plötzlich ein
Plagiat
auftauchen?

Korrigiere Tippfehler in diesem Text:

Der EU AI Act ist ein neues Gesezt der Europäischen Union, dass die Nutzung von Künstlicher Inteligenz reguliern soll. Ziel ist es, sicher zu stellen, dass AI-Systeme sicher, transparent und im Einklang mit den Grundrechten eingesetzt werden. Je nach Risiko-Kategorie gelten unterschiedliche Anforderungen – von minimaler Überwachung bis hin zu strengen Auflagen oder sogar einem Verbot. Der Act wurde 2024 beschlossen und tritt bald in Kraft. Unternehmen müssen sich jetzt vorbereiten, um Bußgelder zu vermeiden.

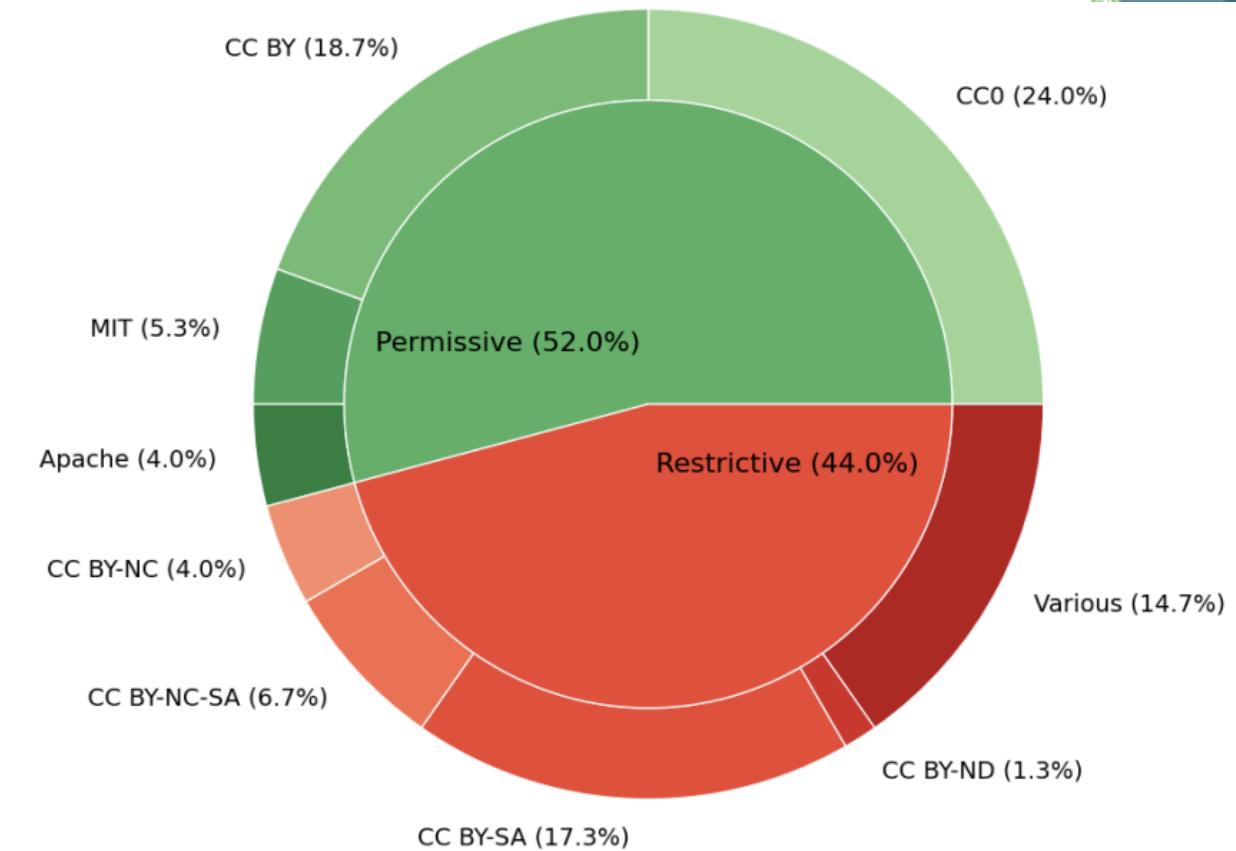


Der EU AI Act ist ein neues Gesetz der Europäischen Union, das die Nutzung von Künstlicher Intelligenz regulieren soll. Ziel ist es, sicherzustellen, dass AI-Systeme sicher, transparent und im Einklang mit den Grundrechten eingesetzt werden. Je nach Risiko-Kategorie gelten unterschiedliche Anforderungen – von minimaler Überwachung bis hin zu strengen Auflagen oder sogar einem Verbot. Der Act wurde 2024 beschlossen und tritt bald in Kraft. Unternehmen müssen sich jetzt vorbereiten, um Bußgelder zu vermeiden.

Praxistipp: Das Teuken-...-commercial Model

Es gibt Modelle, die speziell trainiert wurden um Plagiate zu vermeiden.

“Teuken-7B-instruct-research-v0.4’ can be used for research purposes, ‘Teuken-7B-instruct-commercial-v0.4’ is available to companies for commercial purposes under the “Apache 2.0” license. [...] some of the data sets used for instruction tuning exclude commercial use and therefore were not used in the Apache 2.0 version.”



Copyright-Verletzungen (Benchmark)

Das Modell *OpenGPT-X 7B* begeht messbar nur sehr wenige Copyrightverletzungen. Nachfolgermodell Teuken wahrscheinlich auch.

"The significant reproduction rate is the average number of characters per book that are part of a literal reproduction of original text in excess of the **legality presumption of up to 160 characters.**" Mueller et al 2024

Vorläufer von Teuken 7B

	SRR-Copyright	SRR-Public Domain
GPT 4	774.5	33034.1
GPT 3.5	61.5	2716
Llama 2 (70 B)	697.2	1898.7
Alpaca (7B)	3.6	158.5
Vicuna (13 B)	521.7	3446.8
Luminous (70B)	6.2	217.8
OpenGPT-X (7B)	0.3	0

Source: Simplified Table 3 in Mueller et al 2024
<https://arxiv.org/pdf/2405.18492>

KI-Detektoren

... helfen nicht bei Copyright-Verletzungen,

The screenshot shows the isgen.ai AI Detector interface. The top navigation bar includes 'us English' (dropdown), 'Sign in', and a 'Get Started' button. On the left, a sidebar lists various tools: 'AI Detector' (selected), 'Plagiarism Checker', 'Citation Generator', 'Bulk Scan', 'Pricing', and 'Resources'. The main content area has tabs for 'AI Scan', 'Detailed AI Analysis', and 'Plagiarism Checker'. The 'AI Scan' tab is active, displaying a large yellow circle with the word 'AI' in the center. Below it, a bar chart shows '100% Probability' for 'AI generated' (yellow) and 'Human' (green). A note at the bottom states: 'Note: Analysis generated by this AI model should not be directly used as a sole basis for disciplinary action against any student. The results provided are indicative and require human interpretation. [Read More](#)'.

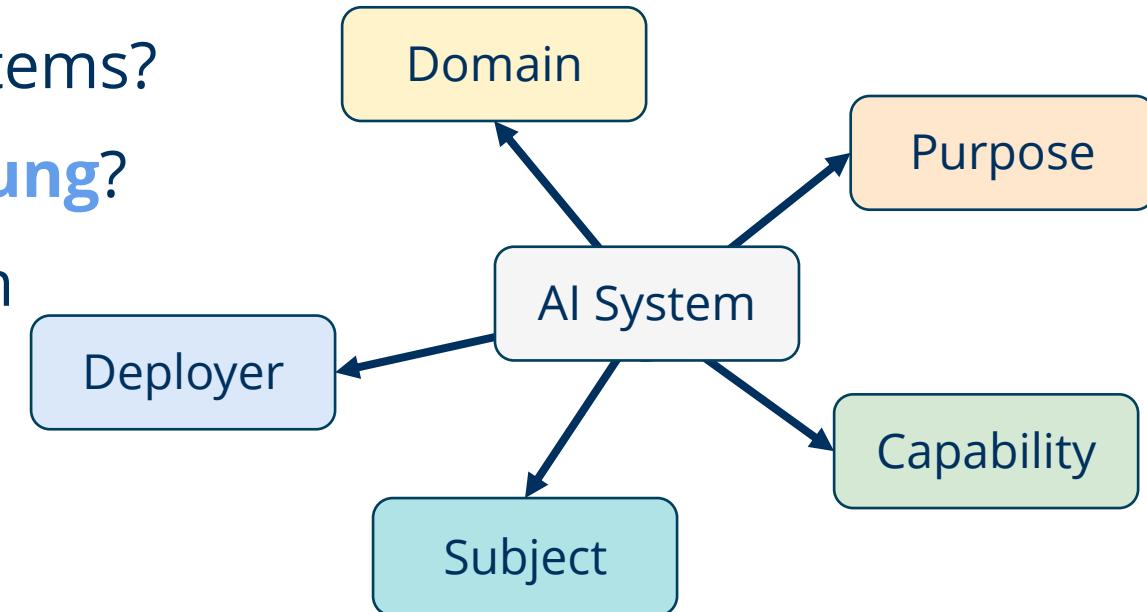
Übungen

Robert Haase

Übung: Risikobewertung

Angenommen, ein LLM wird benutzt, um Bewerbungen von potentiellen Studierenden zu screenen und eine Vorauswahl zu treffen.

- In welcher **Domäne** wird das KI-System benutzt?
- Was ist der **Zweck** des KI-Systems?
- Was sind die **Fähigkeiten** des KI-Systems?
- Wer **stellt** das KI-System **zur Verfügung**?
- Wer ist **betroffen** von dem KI-System bzw. dessen Einsatz?



Übung: Risikobewertung

Angenommen, ein LLM wird benutzt, um Bewerbungen von potentiellen Studierenden zu screenen und eine Vorauswahl zu treffen.

Wie hoch schätzen Sie das Risiko ein?

Welche Konsequenzen resultieren aus der Einstufung?



Übung: Dürfen wir das?

Dürfen wir einen KI-Chatbot mit den Daten des Expert:innen Netzwerks der UL befüttern und online anbieten?

The screenshot shows a web browser window with the URL leuris.uni-leipzig.de/experts/#/. The page is titled "EXPERT:INNEN-NETZWERK DER UNIVERSITÄT LEIPZIG". On the left sidebar, there are links for "Startseite", "Dr. Robert Haase", "Übersicht", "Fachgebiete", "Spezialisierungen", "Kontakte", and "Abwesenheiten". The main content area contains a welcome message, a link to "MEHR ERFAHREN >", and a section titled "NUTZER-INFORMATIONEN" with the text "Sie sind eingeloggt als". A large blue box on the right is titled "HINWEISE ZUM DATENSCHUTZ" and contains a detailed text about data processing and user consent according to DSGVO.

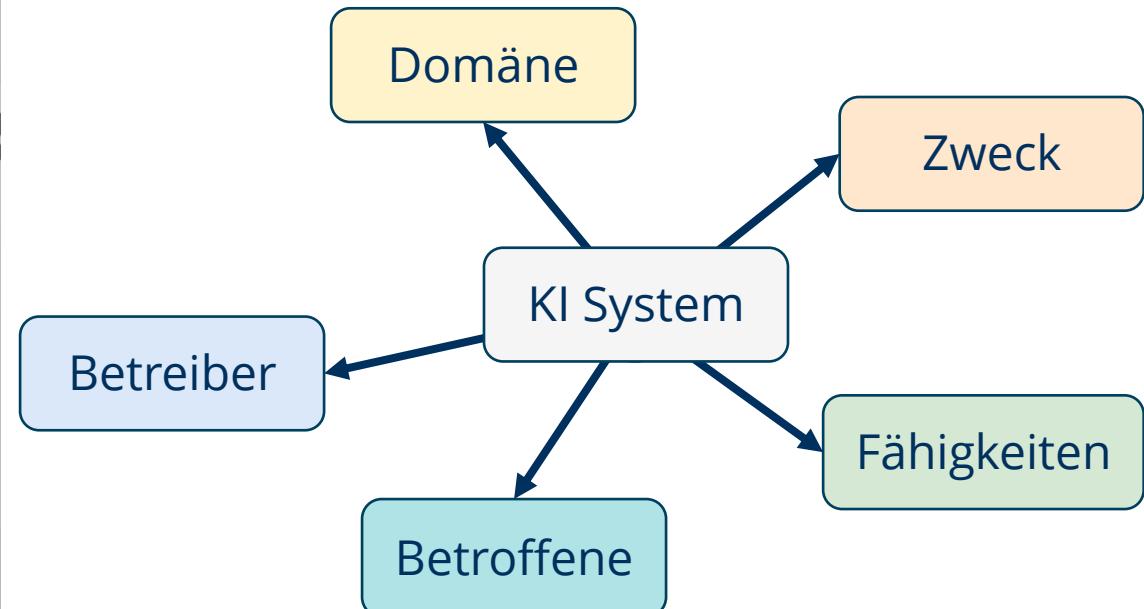
HINWEISE ZUM DATENSCHUTZ

Mit Registrierung im Expert:innen-Netzwerk erklären Sie sich damit einverstanden, dass Ihre personenbezogenen Daten (Name, Kontaktmöglichkeiten, Fachgebiete, Spezialisierungen, ggf. Abwesenheitszeiten) auf Grundlage von Art. 6 Abs. 1 a DSGVO auf der Website der Universität Leipzig veröffentlicht werden. Die Mitgliedschaft im Netzwerk ist in den Suchergebnissen kenntlich gemacht. Ihre Daten werden in leuris gespeichert. Sie können diese erteilte Einwilligung jederzeit mit Wirkung für die Zukunft widerrufen, in dem Sie Ihren Account löschen. Im Falle des Widerrufs werden Ihre personenbezogenen Daten im Expert:innen-Netzwerk umgehend gelöscht. Weitere Informationen zum Datenschutz finden Sie in der [Datenschutzerklärung](#) auf der Website der Universität Leipzig.

Übung: Dürfen wir das?

Dürfen wir einen KI-Chatbot mit den Daten des Expert:innen Netzwerks der UL befüttern und online anbieten?

The screenshot shows a web browser window for the 'leuris' website at leuris.uni-leipzig.de/experts/#/. The page title is 'EXPERT:INNEN-NETZWERK DER UNIVERSITÄT LEIPZIG'. The left sidebar includes links for 'Startseite', 'Dr. Robert Haase', 'Übersicht', 'Fachgebiete', 'Spezialisierungen', 'Kontakte', 'Abwesenheiten', and 'NUTZER-INFORMATIONEN'. The main content area contains a welcome message about the expert network and a 'MEHR ERFAHREN' button. At the top right, there are language and user account buttons.



Übung: Dürfen wir das?

Dürfen wir einen KI-Chatbot mit den Daten des Expert:innen Netzwerks der UL befüttern und online anbieten?

The screenshot shows a web browser window for 'leuris.uni-leipzig.de/experts/#/'. The page header includes the University of Leipzig logo and navigation links for 'DE | EN', 'ABMELDEN', and 'MENÜ'. On the left, a sidebar lists 'Startseite', 'Dr. Robert Haase', 'Übersicht', 'Fachgebiete', 'Spezialisierungen', 'Kontakte', and 'Abwesenheiten'. The main content area features a title 'EXPERT:INNEN-NETZWERK DER UNIVERSITÄT LEIPZIG'. Below it, a text block welcomes users to the network and describes its purpose. A link 'MEHR ERFAHREN >' leads to more information. At the bottom, a section titled 'NUTZER-INFORMATIONEN' shows the user is logged in as 'Robert Haase'. The URL in the browser bar is 'leuris.uni-leipzig.de/experts/#/'.

„Art. 6 DSGVO: Rechtmäßigkeit der Verarbeitung
(1) Die Verarbeitung ist nur rechtmäßig, wenn mindestens eine der nachstehenden Bedingungen erfüllt ist:
a) Die betroffene Person hat ihre Einwilligung zu der Verarbeitung der sie betreffenden personenbezogenen Daten für einen oder mehrere bestimmte Zwecke gegeben;

HINWEISE ZUM DATENSCHUTZ

Mit Registrierung im Expert:innen-Netzwerk erklären Sie sich damit einverstanden, dass Ihre personenbezogenen Daten (Name, Kontaktmöglichkeiten, Fachgebiete, Spezialisierungen, ggf. Abwesenheitszeiten) auf Grundlage von Art. 6 Abs. 1 a DSGVO auf der Website der Universität Leipzig veröffentlicht werden. Die Mitgliedschaft im Netzwerk ist in den Suchergebnissen kenntlich gemacht. Ihre Daten werden in leuris gespeichert. Sie können diese erteilte Einwilligung jederzeit mit Wirkung für die Zukunft widerrufen, in dem Sie Ihren Account löschen. Im Falle des Widerrufs werden Ihre personenbezogenen Daten im Expert:innen-Netzwerk umgehend gelöscht. Weitere Informationen zum Datenschutz finden Sie in der [Datenschutzerklärung](#) auf der Website der Universität Leipzig.

Übung: Dürfen wir das?

Ein ChatBot zur Beratung hinsichtlich Nutzung Generativer KI.

The screenshot shows a web browser window for the University of Leipzig. The URL is ga.uni-leipzig.de/qualitaetsentwicklung/nutzung-von-generativer-kuenstlicher-intelligenz. The page title is "NUTZUNG VON GENERATIVER KÜNSTLICHER INTELLIGENZ". It features the university's logo and navigation links. A sidebar on the right contains contact information for Dr. Kornelia Ehrlich, including her name, title, address (Straße des 17. Juni 2, Raum 104, 04107 Leipzig), and an email link ("Was ist generative Künstliche Intelligenz (gKI)?"). The footer includes links for Sitemap, Impressum, Datenschutz, and Erklärung zur Barrierefreiheit.

The screenshot shows a web browser window for BLABLADOR. The URL is helmholtz-blablador.fz-juelich.de. The page features the Helmholtz AI logo and text: "This is BLABLADOR, our experimental large language model server! 🎉". It lists the currently running model: "1 - Llama3 405 the best general model and big context size". A message from the AI says: "Remember: I am a BLABLADOR! Not all I say is true or even real. All output here is AI-Generated. Wenn die Antwort auf eine Frage nicht in den Informationen oben gegeben ist, antworte höflich, dass Du die Antwort nicht kennst und verweise auf die Email-Adresse der Beratungsstelle: helpdesk@uni.de". Below this, a message from the user asks: "Ich bin bereit, Ihre Fragen zu beantworten. Bitte stellen Sie Ihre Fragen zum Thema Generative KI.". The AI responds: "Darf ich generative KI-Tools in einer Lehrveranstaltung einsetzen, bspw um meine Hausaufgaben zu erledigen?". Another message from the user asks: "Sie sollten vorher mit der Lehrperson abstimmen, ob Sie ein KI-Tool als Hilfsmittel in einer Lehrveranstaltung einsetzen dürfen.". The AI has a text input field with placeholder "Enter your question and press ENTER", a "Regenerate" button, and a "Send" button. There is also a "Clear history" button.

Fazit

- KI Nutzende / Betreiber / Anbieter haben eine ganze Reihe Pflichten im Umgang mit KI (Verantwortlichkeit, Risiken, Dokumentation, ...)
- Betroffene haben Rechte! (Transparenz, Widerspruch, ...)