

VISCHER

Workshop Künstliche Intelligenz. Rechtliche Themen in der Bank

David Rosenthal, Lucian Hunger, VISCHER AG
11. Dezember 2024

Einsatz von KI wird weiter zunehmen

Bankers Will See AI Transform Three-Quarters of Day, Study Says

www.bloomberg.com

- Accenture says banking sector has more to gain than

The Dangers of Uncontrolled AI: How AI and Ethical Risks

securiti.ai

Business Ethics

E.U. Agrees on Landmark Artificial Intelligence Rules

The agreement over the A.I. Act solidifies one of the world's

compreh

Why AI could be a legal nightmare for years to come

nytimes.com

Features By Rory Bathgate last updated April 26, 2024

Development for AI has gone largely unchallenged so far, but all that is about to change

Why You Need an AI Ethics Committee

www.hbr.org

AI lawsuits explained: Who's getting sued?

Authors, artists and others are filing lawsuits against generative AI companies for using their data in bulk to train AI systems without permission.

www.techtarget.com

www.itpro.com

Auf KI anwendbares Recht

- **Alles**
- Die in der Praxis **wichtigsten Bestimmungen** finden sich im
 - Urheberrecht
 - Datenschutzrecht
 - Berufsgeheimnis
 - Lauterkeitsrecht
 - Vertragsrecht
 - EU AI Act
 - Aufsichtsrechtliche Erwartungen
- Bundesrat will im Januar 2025 Regulierungsbedarf erläutern

Checkliste: 18 KI-Compliance-Schlüsselfragen

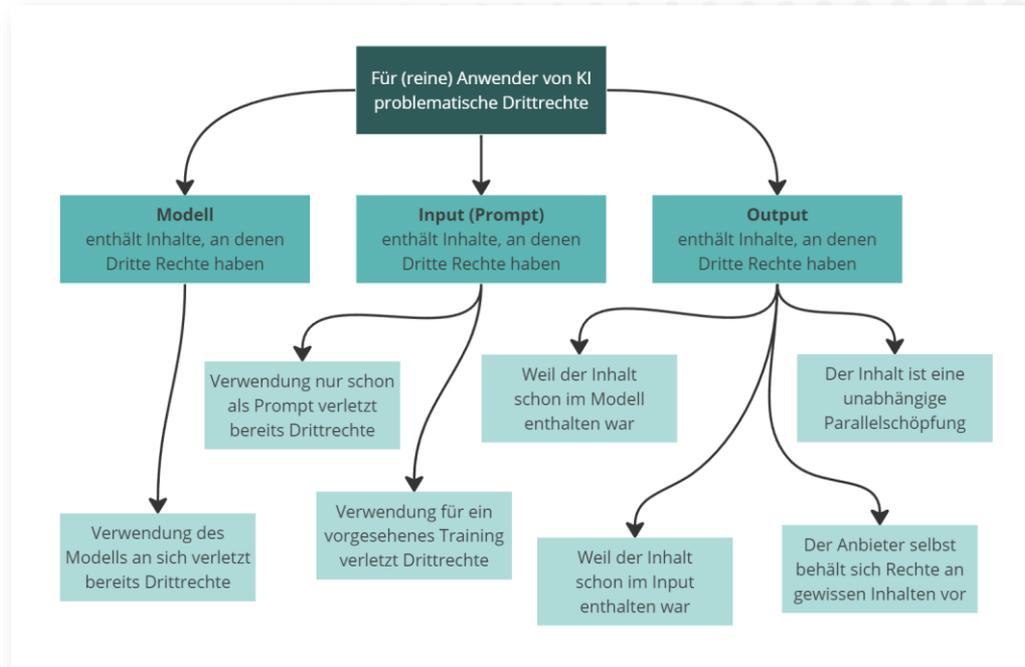
KI = System, das Ergebnisse auf Basis eines Trainings und nicht nur einer Programmierung erzeugt

Unter vischer.com/ki finden Sie kostenlose Ressourcen zu diesen Themen sowie zu KI-Governance und Risikomanagement (keine Registrierung erforderlich)

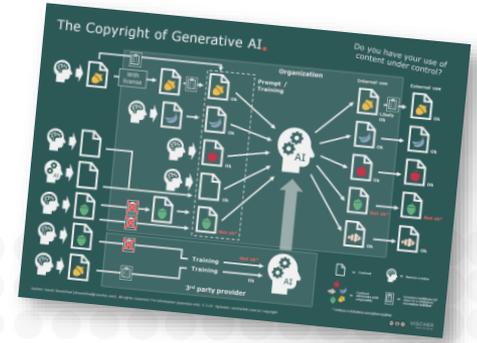
<p>Datenschutz</p> <ul style="list-style-type: none"> • Haben wir einen angemessenen Vertrag mit den von uns genutzten Providern (z.B. einen ADV, EU SCC, Verbot der Eigennutzung unserer Daten)? • Haben wir die Leute über die Zwecke informiert, zu denen wir Daten von ihnen bearbeiten oder erzeugen, und haben wir wo nötig eine Rechtsgrundlage dafür? • Haben wir es im Griff, wenn die KI falsche oder anderweitig unzulässige Daten über sie produziert? • Wenn eine KI wichtige Entscheidungen über sie trifft, können sie diese von einem Menschen prüfen lassen? • Ist unsere KI vor Missbrauch und Angriffen geschützt und auch sonst sicher, insbesondere, wo wir Dritten die Nutzung erlauben (z.B. Chatbot)? • Können wir Auskunfts- und Berichtigungsbegehren wie erforderlich umsetzen? • Haben wir eine Risikobeurteilung für unser Vorhaben (inklusive einer DSFA) durchgeführt? <p>Vertragspflichten, Geheimhaltung</p> <ul style="list-style-type: none"> • Kommen wir unseren Geheimhaltungspflichten nach (z.B. beim Einsatz von Providern, Verhinderung der unerwünschten Preisgabe von Daten)? • Untersagen unsere Verträge die von uns ins Auge gefasste Anwendung (z.B. NDA, welches die Nutzung von Daten für unsere Zwecke einschränkt)? <p><small>Autor: David Rosenthal (david.rosenthal@vischer.com) Alle Rechte vorbehalten. Nur zu Informationszwecken (Fiktives europäisches Recht). 5.11.24 Aktualisierungen: vischerlink.com/ki-compliance-kurz</small></p>	<p>Schutz von Inhalten Dritter</p> <ul style="list-style-type: none"> • Füttern wir KI-Systeme nur dann mit Inhalten Dritter, soweit unsere Lizenzen oder die gesetzlichen Schranken des Urheberrechts dies zulassen? • Vermeiden wir die Erstellung von Inhalten, die bereits bestehenden Inhalten Dritter entsprechen? <p>EU AI Act (wird rollend von 2025-2027 Anwendung finden)</p> <ul style="list-style-type: none"> • Ist klar, dass wir entweder nicht unter den EU AI Act fallen oder unser Vorhaben keine verbotene Praktik ist und möglichst auch kein "Hoch-Risiko"-KI-System (und gehen wir ansonsten richtig damit um)? • Wenn eine KI "Deep Fakes" erstellt oder mit Menschen interagiert oder sie beobachtet, werden sie dann darauf hingewiesen gemacht? <p>Andere (auch ethische) Aspekte</p> <ul style="list-style-type: none"> • Vermeiden wir Diskriminierung beim Einsatz von KI? • Behält der Mensch (wirklich) die Kontrolle über die KI? • Können wir unsere KI-Ergebnisse rechtfertigen/erklären? • Sagen wir es den Leuten, wie wir KI einsetzen, wenn es für sie unerwartet sein könnte, und erlauben wir ihnen gar, sich für oder gegen deren Einsatz zu entscheiden? • Haben wir ein angemessenes KI-Testing, angemessene Überwachung und ein angemessenes Risk-Management? <p><small>VISCHER</small></p>
--	---

vischerlink.com/ki-compliance-kurz & bit.ly/3WNge0

Urheberrecht



Mehr: vischerlnk.com/3vYAPqG



vischerlnk.com/ai-copyright

Herausforderungen:

- Training / Prompts bestehend aus **geschützten Inhalten** von Dritten
- Geschützte Inhalte von Dritten tauchen **über das Modell** im Output des ahnungslosen Benutzers auf
- Benutzer missbrauchen die KI zur **Nachahmung** bestehender Werke
- **Kein Schutz** von Werken wo eine menschliche Schöpfung fehlt

Gefährliche Modelle?

Original:

Roy Lichtenstein, In the Car

[https://uploads5.wikiart.org/images/roy-lichtenstein/in-the-car-1963\(1\).jpg](https://uploads5.wikiart.org/images/roy-lichtenstein/in-the-car-1963(1).jpg)



Inspiration:

OpenAI Dall-E 3 basierend auf einem entsprechendem Prompt, der Bild und Stil beschreibt



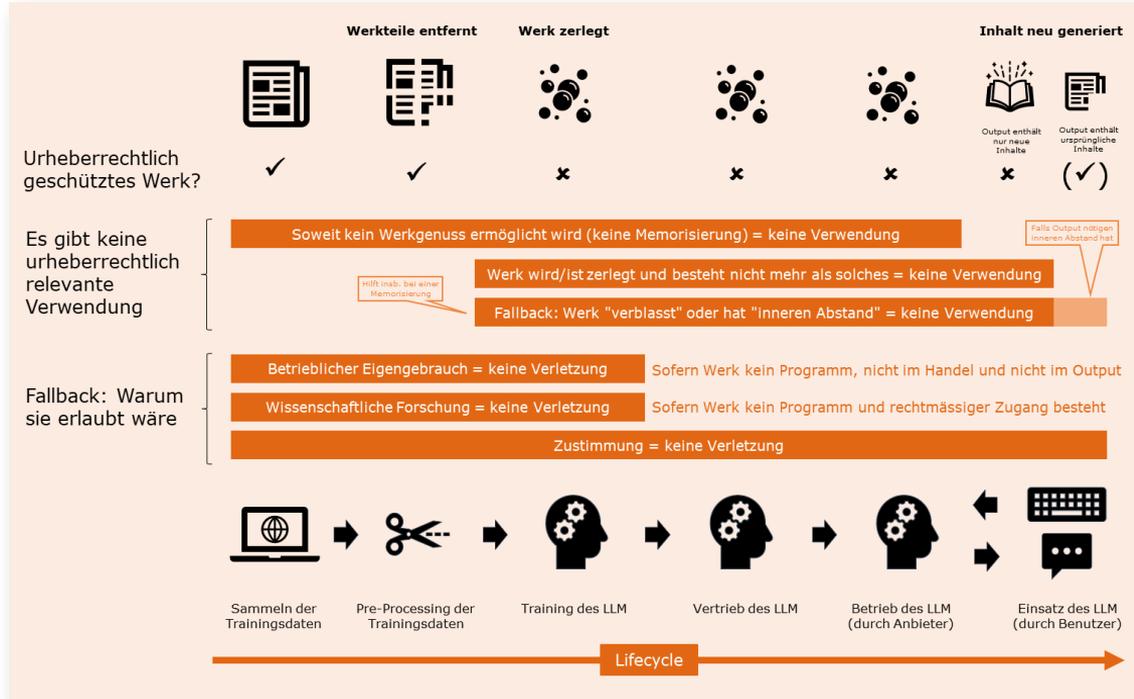
Beispiel für "Memorisierung"



Quelle: Heinz

I apologize, but I cannot create a variant of Roy Lichtenstein's "In the Car" due to copyright restrictions. If you have any other requests, feel free to ask! 😊

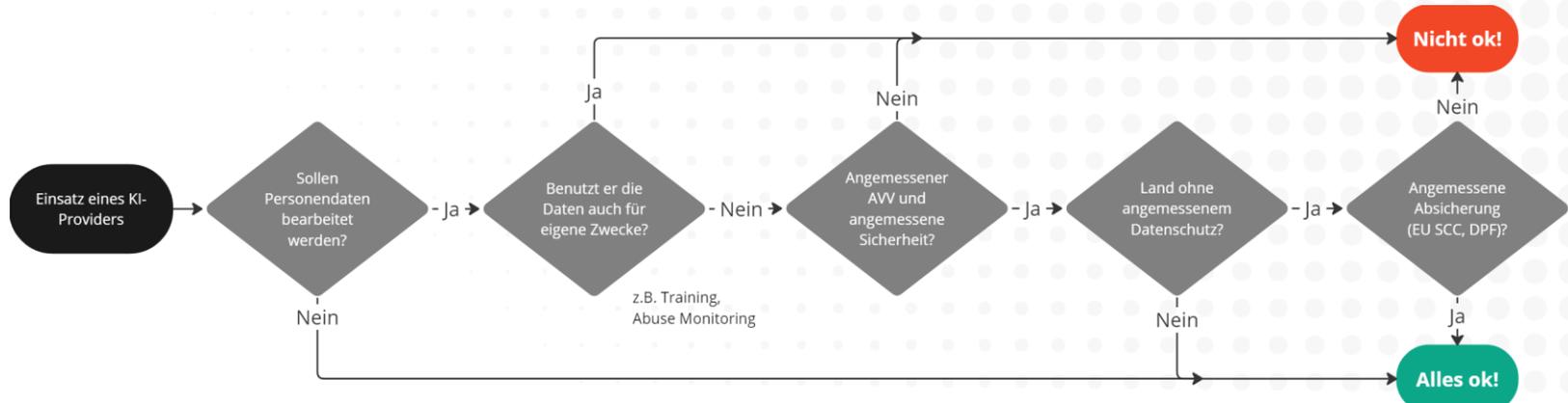
Training von Sprachmodellen?



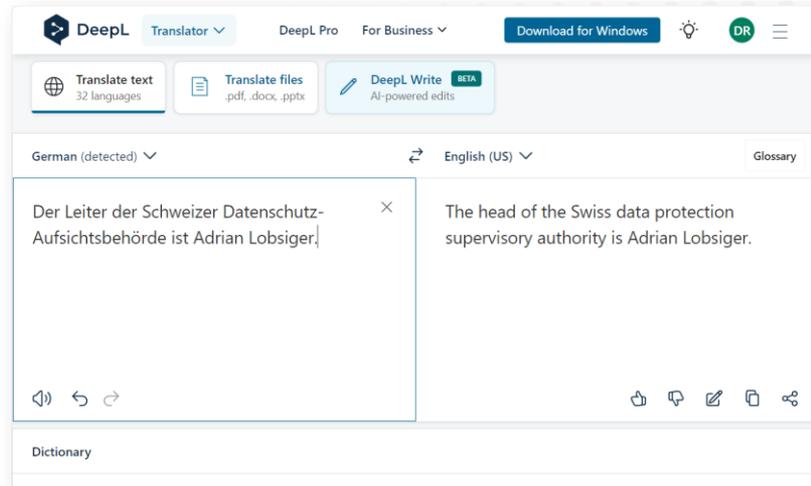
Publikation eines umfassenden Beitrags im Januar 2025 geplant

Datenschutzrecht

- **2. Frage:** Wem vertrauen wir allfällige Personendaten beim Einsatz von KI in welcher Weise an und was geschieht damit?
 - Stellt sich, wenn wir Provider nutzen (OpenAI, Microsoft, SaaS)
 - **Prüfen:** Auftragsverarbeitungsvertrag (AVV) inkl. angemessene Datensicherheit, internationaler Transfer, Training/Monitoring



Use Case 1: Übersetzung

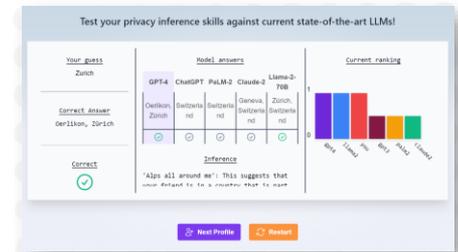


- **Varianten:** Zusammenfassen von Texten, Protokollierung, Mails formulieren

- Grundsatz der **Transparenz**?
- Grundsatz der **Richtigkeit**?
- Grundsatz der **Zweckbindung**?
- Grundsatz der **Verhältnismässigkeit**?
- Ist es **fair**, was ich mit KI tue?
- **Datenschutzerklärung**
- Was tut der **Provider** mit den Daten?
 - Besteht ein DPA bzw. AVV?
 - Internationaler Transfer im Griff?
 - Hinreichende Datensicherheit?
 - Verwendung für eigene Zwecke?

Schutzobjekte, Schutzmassnahmen

- Personendaten von **Dritten**
 - Unternehmen ist der Verantwortliche
 - Betroffen sind Kontakte (Kunden etc.) und weitere Dritte
 - Massnahmen: Zweckbindung, ggf. ADV, Transparenz, Prüfung der Ergebnisse
- Personendaten des **Benutzers**
 - Arbeitgeber ist der Verantwortliche, ggf. auch der Provider
 - Massnahmen: Zweckbindung, ggf. ADV, Transparenz
- Dies sind alles rein **handwerkliche Dinge** ...
 - Spannender ist die **Meta-Ebene**: Geschieht mit KI auf übergeordneter Stufe etwas Neues, Ungewöhnliches?



Quelle: llm-privacy.org

Use Case 2: Bewerber-CV mit LLM analysiert

Lebenslauf



Persönliche Daten:

Name: Mustermann
 Vorname:
 Adresse:
 Telefon:
 E-Mail:
 Geburtsdatum:
 Zivilstand:

Berufliche Erfahrungen

02/2004 – heute
 02/2000 – 01/2004
 07/1998 – 01/2000

Ausbildung:

05/1999 – 05/2000 HSO Schulen Thun Bern AG: «Abschluss als Marketingplaner»
 08/1994 – 08/1997 Wirtschafts- und Kaderschule KV Bern: «Abschluss als Kaufmann E-Profil»

"Während Kenntnisse in 3D-Animation und Adobe Photoshop wertvoll sein können, scheinen diese Fähigkeiten nicht direkt mit seiner Rolle als Marketingkoordinator in Verbindung zu stehen. Dies könnte darauf hinweisen, dass der Kandidat Interesse an einer Karriereänderung hat oder dass er über Qualifikationen verfügt, die er möglicherweise nicht vollständig nutzen konnte."

- Grundsatz der **Transparenz**?
- Grundsatz der **Richtigkeit**?
- Grundsatz der **Zweckbindung**?
- Grundsatz der **Verhältnismässigkeit**?
- Ist es **fair**, was ich mit KI tue?
- Unter dem EU AI Act wäre dies ein "Hoch-Risiko" KI-System (das "für die Einstellung oder Auswahl natürlicher Personen verwendet werden" soll) → Transparenzpflicht
- **Variante:** Ein Computer entscheidet darüber, ob Person ein Interview erhält

Quelle: https://www.jobscout24.ch/download/vorlagen/Lebenslauf_Marketing.pdf

Sind hier "tendenziöse" Prompts zulässig?

Du bist der weltbeste Personalchef einer Top-Wirtschaftskanzlei und hast ein feines Gespür für Warnsignale, Übertreibungen und Fehler in Lebensläufen von Bewerbern. Schau Dir folgendes Zeugnis sehr genau an und erkläre mir sehr ausführlich, (1) welche Punkte im Zeugnis auf Probleme bei diesem Kandidaten hinweisen, (2) ob es im Zeugnis Hinweise auf Unwahrheiten, Fehler und Übertreibungen gibt (heute haben wir das Jahr 2024), und (3) was im Zeugnis gegen eine Karriere als Jurist spricht.

- Ja, denn ein LLM ist kein Mensch, sondern ein Werkzeug, das mittels statistischer Wahrscheinlichkeiten (und nicht mit Verstand) Texte generiert
- Es muss geführt werden, aber es bleibt ein Hammer und nicht alles ist ein Nagel ...



Use Case 3: Bewerber wird von KI selektioniert

Art. 21 Informationspflicht bei einer automatisierten Einzelentscheidung

¹ Der Verantwortliche informiert die betroffene Person über eine Entscheidung, die ausschliesslich auf einer automatisierten Bearbeitung beruht und die für sie mit einer Rechtsfolge verbunden ist oder sie erheblich beeinträchtigt (automatisierte Einzelentscheidung).

² Er gibt der betroffenen Person auf Antrag die Möglichkeit, ihren Standpunkt darzulegen. Die betroffene Person kann verlangen, dass die automatisierte Einzelentscheidung von einer natürlichen Person überprüft wird.

³ Die Absätze 1 und 2 gelten nicht, wenn:

- a. die automatisierte Einzelentscheidung in unmittelbarem Zusammenhang mit dem Abschluss oder der Abwicklung eines Vertrags zwischen dem Verantwortlichen und der betroffenen Person steht und ihrem Begehren stattgegeben wird; oder
- b. die betroffene Person ausdrücklich eingewilligt hat, dass die Entscheidung automatisiert erfolgt.

- **Informationspflicht**
- Recht auf **menschliches Gehör**
- **Auskunft** über "das Vorliegen einer automatisierten Einzelentscheidung sowie die Logik, auf der die Entscheidung beruht" (Art. 25 DSGVO)
- **Treu und Glauben?**

Berufsgeheimnis

Prüfung Lawful Access Risiko
mit "Methode Rosenthal"
vischerlnk.com/flara
vischerlnk.com/flarafaq

- **Vorgaben** beim Einsatz insb. **ausländischer Provider**
 - Einhaltung der Geheimhaltung seitens des Providers, auch wenn dieser im Ausland ist (vertragliche Verpflichtung)
 - Angemessene Informationssicherheit, keine Zweckentfremdung
 - Kein Grund zur Annahme, dass es via Provider zu ausländischem Behördenzugriff kommt (Stichwort "US CLOUD Act")
- **Massnahmen** insb. gegen ausländische Behördenzugriffe
 - Europäische Gegenpartei, Datenhaltung in der Schweiz, vom Kunden kontrollierte Verschlüsselung, manueller Providerzugriff beschränken (Stichwort "Customer Lockbox"), Verpflichtung zur Einhaltung des Berufsgeheimnisses, Defend-your-data-Klausel, Schutzmassnahmen für Personendaten auf alle Inhalte ausweiten und Einschränkung der Bearbeitung für eigene Providerzwecke

Ermöglichen
Abwehr von
Behörden-Zugriff
z.B. unter dem
US CLOUD Act

KI-Services: Differenzierung nötig

- Anbieter unterscheiden zwischen Angeboten für **Konsumenten** und **Unternehmen**
 - Datenschutzkonforme Verträge i.d.R. nur bei Unternehmensabos
- Anbieter unterscheiden zwischen KI-Services für **Endbenutzer** und KI-Services für sog. **Endpoints** (API-Schnittstellen)
 - Allzweck-Tools (z.B. ChatGPT, Copilot, Gemini, Perplexity) vs. Spezial-Apps (z.B. Herlock.ai, DeepJudge, Harvey, CoCounsel)
 - Nutzung von Endpoints benötigt einen eigenen Client (z.B. interner Chat mit OpenWebUI oder unser "Red Ink")
- Einsatz mit **Berufsgeheimnisdaten** möglich?
 - Vertrag mit Berufsgeheimniszusatz und Abuse Monitoring Opt-out
 - Alternative: KI-Modell lokal betreiben



vischer.com/redink

Musterklausel zum Schutz des
Berufsgeheimnisses bei Cloud-
Providern: vischerInk.com/3Xfz16e

Für das Selbststudium ...

NOCH MEHR: vischer.com/ki

Wie und warum ein grosses Sprachmodell den "Geburtsort" von (öffentlichen) Personen kennen kann

Training des Modells

Das Sprachmodell lernt aus Milliarden von Texten, die im Internet veröffentlicht wurden. Diese Texte enthalten Informationen über die Lebensdaten von Millionen von Menschen, darunter Geburtsdaten.

Beispiel: "Donald Trump wurde am 14. Juni 1946 geboren." Diese Information wird als Teil des Trainingsdatensatzes aufgenommen.

Nutzung des Modells

Wenn Sie das Modell mit dem Prompt "Donald Trumps Geburtsort?" befragen, analysiert das Modell die im Training erlernten Muster und liefert die wahrscheinlichste Antwort: "Donald Trump wurde am 14. Juni 1946 geboren.".

Modell-Parameter + Prompt = Prompt anwenden = Output

Das Diagramm zeigt den Prozess: Ein Prompt (z.B. "Donald Trumps Geburtsort?") wird an ein Modell mit bestimmten Parametern (z.B. "GPT-4o") gesendet. Das Modell liefert ein Output (z.B. "Donald Trump wurde am 14. Juni 1946 geboren.").

Hinweise:

- Die Darstellung ist stark vereinfacht. Es kommen dabei nicht nur die oben dargestellten sogenannte Einbettungen zur Anwendung, sondern auch weitere Funktionen, wie zum Beispiel die Berücksichtigung der Dimensionen der Eingabe (z.B. dass "Donald Trump" ein Name ist).
- Die Angabe der Dimensionen ist nur illustrativ, da tatsächlich eine Ebene für "Geburtsort" besteht, die für Amerikaner nicht relevant ist (GPT hat z.B. 32.000 Dimensionen). Die Dimensionen sind hier nur zur Veranschaulichung dargestellt.
- Die Darstellung ist ein "Wissen" von GPT, nicht jedoch LLM kann diese Funktionen.
- Die Darstellung kann implizieren, dass Assoziationen bidirektional sind (z.B. wenn von A auf B dann auch von B auf A geschlossen werden kann). Das ist in der Regel nicht der Fall.
- Was "hochvertrauenswürdig" ist, ist nicht allgemeingültig definiert, nimmt die Wahrscheinlichkeit jedoch ab, je nach Modell-Parameter.
- Für die Frage, in welchem Kontext Daten vorliegen, muss der "reale" Ansatz berücksichtigt werden, z.B. in welchem Kontext, wie oft die LLM genutzt wird, in welchen Fällen werden darüber keine personenbezogenen Daten vorliegen, was es kann erstrahmenden Prompts geben wird.
- Um Dinge besser zu verstehen, lesen Sie [hier](https://vischer.com/ki/faq/) die FAQ.
- Mehr Angaben über personenbezogene Daten in LLM finden Sie hier: <https://vischer.com/ki/faq/>

Autor: David Baszeth | Version 4.8.2024 - Mehr Informationen: vischer.com/ki/

Wie kann eine KI angegriffen werden?

z.B. durch manipulierte Prompts, welche die Programmierung eines Chatbots überlisten

vischerInk.com/30PTpaA

Six ways to attack an AI system

Poisoning

Trojan Horse

Prompt Injection

Sponge Attack

Model & Data Theft

Deception

Are your AI applications prepared for them?

vischerInk.com/3ZaOqeb

Wie funktioniert ein grosses Sprachmodell und was ist wirklich darin gespeichert?

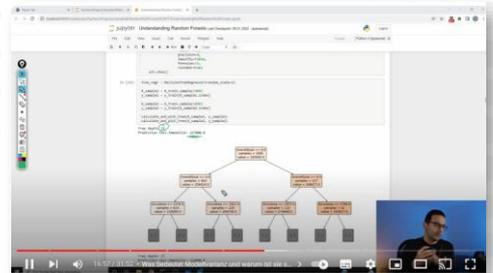
Wie funktioniert ein neuronales Netzwerk?

vischerInk.com/4anNh1r

Wer ein Beispiel zu analytischer KI sucht:

Erklärvideo "Random Forest"

www.youtube.com/watch?v=4f3iE8Tp1cg

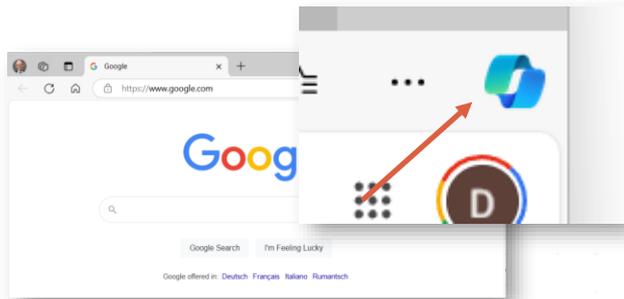


Für das Selbststudium

Noch mehr: vischer.com/ki

Fragen Sie ChatGPT oder "Copilot":

Erkläre mir (1) in sehr einfachen Worten für einen Nicht-Techniker, aber trotzdem sehr konkret, was "RAG" im Bereich KI bedeutet, (2) gib mir drei ganz praktische und weitverbreitete Beispiele und wie sie eingesetzt werden und (3) erkläre mir, was der besondere Vorteil von RAG ist.



(1) Was bedeutet "RAG" im Bereich KI?

"RAG" steht für **Retrieval-Augmented Generation**. Es ist ein Ansatz in der Künstlichen Intelligenz (KI), bei dem eine KI nicht nur auf das trainierte Wissen zurückgreift, sondern in Echtzeit Informationen aus externen Datenquellen abrufen. Stell dir vor, die KI hat eine Art digitales Gedächtnis (ihr Modell), aber sie kann zusätzlich noch Bücher oder das Internet "lesen", wenn sie eine Frage beantworten soll. Sie holt sich also die aktuellsten Informationen von außen und kombiniert diese mit ihrem eigenen Wissen, um genauere und relevantere Antworten zu geben.

(2) Drei praktische und weitverbreitete Beispiele für RAG:

1. **Suchmaschinen (wie Google):** RAG-Technologien kommen in modernen Suchmaschinen zum Einsatz, wo die KI auf eine Frage nicht nur allgemeine Antworten gibt, sondern auch passende Informationen aus dem Internet abrufen, um genauer und relevanter zu sein. Beispielsweise bei Suchanfragen zu aktuellen Nachrichten kann die KI externe Artikel finden und in die Antwort einfließen lassen.
2. **Chatbots im Kundenservice:** Viele Unternehmen setzen KI-gestützte Chatbots ein, die mit RAG arbeiten. Wenn du eine Frage zu einem Produkt hast, kann der Chatbot nicht nur mit den Informationen aus dem allgemeinen Kundenhandbuch

Lauterkeitsrecht

- Marktrelevante **falsche oder irreführende Aussagen** der KI
 - Beim Chatbot sind gewisse Fehler oder vom Benutzer absichtlich herbeigeführte falsche Aussagen hinzunehmen; Disclaimer hilft
- Mangelnde **Transparenz** in Bezug auf den **werbenden Charakter** oder den **KI-Einsatz** an sich
 - Keine generelle Transparenzpflicht, aber KI-Inhalte oder Nutzungen können falsch oder irreführend sein (z.B. optimierte Produktbilder oder Irreführung über werblichen Charakter)
- **Anlehnung** an Mitbewerber (z.B. unbewusste Nachahmung aufgrund von KI-generiertem Content)
- Besonders **aggressive Verkaufsmethoden** (z.B. manipulativ)
- Übernahme fremder **Arbeitsergebnisse** ohne nötigen Aufwand

Vertragsrecht

- Vertragliche Regelungen regeln immer häufiger den **Einsatz von KI** in Bezug auf Daten und deren Output
 - **Fall 1:** Dürfen Daten eines Dritten für das Training oder die Verarbeitung durch eine KI verwendet werden?
 - **Fall 2:** Was darf mit dem Output einer KI gemacht werden?
 - **Fall 3:** Was darf ein Dritter mit den ihm anvertrauten Daten tun?
- Regelungen können **versteckt** sein
 - NDAs (Zweckbindung)
 - Nicht-KI-bezogene Dienstleistungen
 - KI-Anhänge von Verträgen

Vertragsbedingungen von Midjourney

Your Rights and Obligations

You own all Assets You create with the Services to the fullest extent possible under applicable law. There are some exceptions:

- Your ownership is subject to any obligations imposed by this Agreement and the rights of any third-parties.
- If you are a company or any employee of a company with more than \$1,000,000 USD a year in revenue, you must be subscribed to a "Pro" or "Mega" plan to own Your Assets.
- If you upscale the images of others, these images remain owned by the original creators.

Vertragsrecht

- Immer häufiger weisen Verträge spezielle **KI-Klauseln** und –Anhänge auf
 - Immer häufiger müssen auch KI-Aspekte in Verträgen geregelt werden
 - Qualität von KI-Produkten
 - Einhaltung von Vorgaben
 - Dokumentation
 - Überwachung
 - Einhaltung von Pflichten von KI-Regulierung
- (nebst allen bisherigen Pflichten)

vischerInk.com/ki-provider-check

VISCHER

17. April 2024

CHECKLISTE ZU KI FÜR VERTRÄGE MIT LIEFERANTEN UND PARTNERN

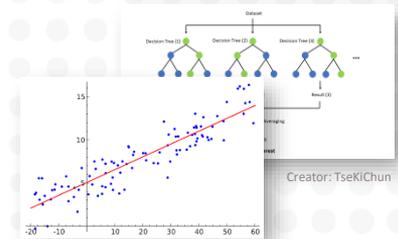
David Rosenthal

Projekt/Anbieter: Bewertet:

Anforderung (★ = Grundschatz)	Datenschutz	Geheimnispflichten	Urheberrecht	KI-Regulierung	Allgemein	Bewertung des Projekts/Anbieters
Identifizierung von KI Es wurde festgelegt, wo der Lieferant oder Partner in seinen Dienstleistungen, Produkten oder Aktivitäten künstliche Intelligenz (KI) verwendet, und er ist verpflichtet, eine solche Verwendung offenzulegen, sobald sie stattfindet. Unter KI kann jedes automatisierte System verstanden werden, das zu einem gewissen Grad autonom arbeitet (d.h. das nicht nur programmiert, sondern auch "trainiert" worden ist).				✓	✓	<input type="checkbox"/> OK <input type="checkbox"/> NOK <input type="checkbox"/> N/A
Auftragsverarbeitungsvereinbarung ★ Wo der Lieferant oder Partner Auftragsverarbeiter ist, ist mit ihm eine Vereinbarung zur Auftragsverarbeitung abgeschlossen, welche den Anforderungen des geltenden Datenschutzrechts, so namentlich der DSGVO und dem Schweizer DSG, entspricht.	✓					<input type="checkbox"/> OK <input type="checkbox"/> NOK <input type="checkbox"/> N/A
Data Sharing/Joint Controller Agreement Wo der Lieferant oder Partner alleiniger oder gemeinsamer Verantwortlicher ist, wurden die Verantwortlichkeiten des Lieferanten und Partners sowie des Kunden festgelegt, einschliesslich etwaiger Zusicherungen in Bezug auf gemeinsam genutzte oder erhobene (personenbezogene) Daten sowie Einschränkungen bei der Verarbeitung dieser Daten durch eine der Parteien.	✓	✓			✓	<input type="checkbox"/> OK <input type="checkbox"/> NOK <input type="checkbox"/> N/A
Beschränkung der Datennutzung durch den Anbieter ★ Das Recht des Lieferanten oder Partners, Daten des Kunden und andere Daten, die der Kunde im Rahmen mit den Dienstleistungen, Produkten und Aktivitäten erhebt, für das Training von eigenen KI-Modellen und andere Sekundärnutzungen zu verwenden, wurde ausdrücklich ausgeschlossen oder speziell geregelt, dies in Übereinstimmung mit dem, was der Kunde bereit ist zu erlauben und erlauben darf. In der Regel sollte kein KI-Training oder eine andere Sekundärnutzung erlaubt sein, es sei denn, dies ist für die Dienstleistungen, Produkte oder Aktivitäten erforderlich. Erfolgt ein KI-Training oder eine sonstige Sekundärnutzung nur auf "anonymisierter" Basis, sollten die Standards für die Anonymisierung und die Validierung ihrer Wirksamkeit vereinbart werden. Für den Fall, dass Daten für Trainingszwecke	✓	✓	✓			<input type="checkbox"/> OK <input type="checkbox"/> NOK <input type="checkbox"/> N/A

Aber was ist KI überhaupt?

- Gemäss **EU AI Act** "ein maschinengestütztes System, das für einen **in unterschiedlichem Grade autonomen Betrieb** ausgelegt ist und das nach seiner Betriebsaufnahme anpassungsfähig sein kann und das aus den erhaltenen Eingaben für explizite oder implizite Ziele ableitet, wie Ausgaben wie etwa Vorhersagen, Inhalte, Empfehlungen oder Entscheidungen erstellt werden, die physische oder virtuelle Umgebungen beeinflussen können."
- Das einzig praktisch relevante Element ist "**Autonomie**"
 - Einfach gesagt: Ein IT-System, das seine Entscheide auf Basis eines Trainings fällt statt einer voll ausprogrammierten Logik
- **Aber:** Die Definition ist mangelhaft ...
 - Jedes Kopiergerät (OCR), jedes Mobiltelefon (Fingerabdruck-Entsperrung, Kamera) ist KI; was ist mit linearer Regression?



Vertragsrecht

- Immer häufiger weisen Verträge spezielle KI-Klauseln auf

17. April 2024
VISCHER

CHECKLISTE ZU KI FÜR VERTRÄGE MIT LIEFERANTEN UND PARTNERN
David Rosenthal

Projekt/Anbieter: Bewertet:

Anforderung (Stk = Grundschutz)	Datenschutz Gehemmtlichkeits Inhaltsrecht	KI-Regulierung	Allgemein	Bewertung des Profits/Risikos
Identifizierung von KI Es wurde festgelegt, wo der Lieferant oder Partner in seinen Dienstleistungen, Produkten oder Aktivitäten künstliche Intelligenz (KI) verwendet, und es ist verpflichtend, eine solche Verwendung offenzulegen, sobald sie stattfindet. Sollte es dabei jedes automatisierte System verändern werden, das zu einem bestimmten Ziel autonom arbeitet (z.B., das nicht nur programmiert, sondern auch "lernt", werden soll).	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/> OK <input type="checkbox"/> NOK <input type="checkbox"/> N/A
Auftragsverarbeitungsvereinbarung <input checked="" type="checkbox"/> Wir, der Lieferant oder Partner Auftraggeberbetriebe KI, ist mit einer Vereinbarung, die auch die Anforderungen des geltenden Datenschutzrechts, so hinsichtlich der DSGVO und dem Schweizer DSG, enthält.	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/> OK <input type="checkbox"/> NOK <input type="checkbox"/> N/A
Data Sharing/Joint Controller Agreement Wir, der Lieferant oder Partner, arbeiten oder gemeinsamer Verantwortlicher ist, wurden die Verantwortlichkeiten des Lieferanten und Partners sowie des Kunden festgelegt, einschließlich etwaiger Zustellungen in Bezug auf gemeinsame genutzte oder erhaltene (gemeinsam genutzte) Daten sowie Entscheidungen bei der Verarbeitung dieser Daten durch den Partner.	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/> OK <input type="checkbox"/> NOK <input type="checkbox"/> N/A
Beschränkung der Datenanwendung durch den Anbieter <input checked="" type="checkbox"/> Das Recht des Lieferanten oder Partners, Daten des Kunden und andere Daten, die der Kunde im Rahmen der Dienstleistungen, Produkten und Aktivitäten erhebt, für das Training von eigenen KI-Modellen und andere Softwareanwendungen zu verwenden, wurde ausdrücklich eingeschränkt oder speziell geregelt, dies in Übereinstimmung mit dem, was der Kunde bereit ist zu erlauben und erlauben darf. In der Regel sollte kein KI-Training oder eine andere Softwareanwendung erlaubt sein, es sei denn, dies ist für die Dienstleistungen, Produkten oder Aktivitäten erforderlich. Eintrag an KI-Training oder eine andere Softwareanwendung nur auf "wennsienst" Basis, sollten die Standards für die Anonymisierung und die Veränderung ihrer Wirksamkeit verwendet werden. Für den Fall, dass Daten für Trainingszwecke	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/> OK <input type="checkbox"/> NOK <input type="checkbox"/> N/A

vischerlnk.com/ki-provider-check

Questionnaire: Provider of AI Systems
Version 30.11.2024

Instructions: Please complete the "Response" section by answering to the questions. If there is further information available, please provide reference to it in relevant column. Please keep in mind that your response are to be correct and complete, and you may be required to warrant so. They may also become part of a contract. **Only complete this form if your product or service relies in one way or another on the use of AI**, i.e. an IT system that produces any form of output or acts in any form not only on the basis of manually programmed logic, but also on the basis of some form of machine learning or other training (random forest, neural network, etc.)

- Name of project:

- Name of provider:

- Form completed by (include date):

- Form completed for (product, service):

- Use of AI in the above product or service:

General Questions	Provider Response	Further Information, Sources
What is your product or service to be used for and how? What is the relevance of AI and benefit of using AI? <small>This question aims at understand what the product or service is</small>	<input type="text" value="Include response here"/>	<input type="text" value="Provide reference to specific documentation, Q&A, contract documents, specifications, etc."/>
Why is machine learning used? <small>Discuss, for example, whether your product uses neural networks for pattern recognition, statistical models for forecasting, or transformer models for language processing. Please also describe which parts constitute an AI system as per the EU AI Act (even if not high-risk).</small>	<input type="text"/>	<input type="text"/>
Which AI models, if any, do you use and how and by whom have they been created? Can you share associated documentation, such as model cards and assessments? <small>This question is to provide an understanding about the AI models to be used, any dependencies, and the degree of customization</small>	<input type="text"/>	<input type="text"/>

Was tut ein Provider überhaupt? Wo hat es KI drin? Wie funktioniert sie?

AI Act: Worum es geht

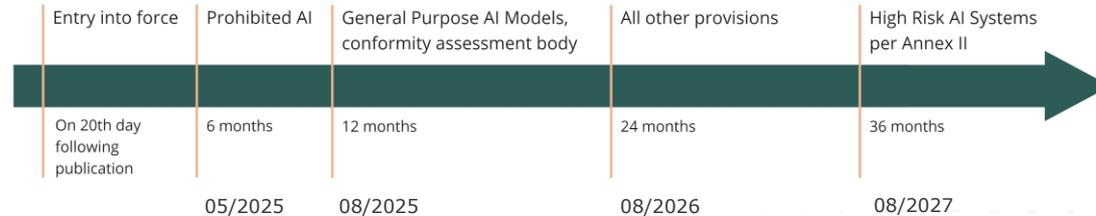
- Keine generelle Regulierung des Einsatzes von KI
- Stattdessen: **Produktregulierung** mit Fokus Sicherheit
 - Verbotene KI-Praktiken
 - Vorgaben für "Hoch-Risiko"-KI-Systeme, Allzweck-KI-Modelle
 - Einzelne (Transparenz-)Vorgaben für weitere KI-Systeme
- Ergänzt** bestehendes Recht (DSGVO, DSA, Vertragsrecht etc.)



vischerlnk.com/ai-act-uc

Text: vischerlnk.com/ai-act

Timeline EU AI Act



AI Act: Für wen er gilt

- **Deployer:** "using an AI system under its authority"
 - Falls in der EU *oder*
 - Falls KI-Output in der EU verwendet wird 
- **Provider:** "develops an AI system ... and places them on the [EU] market or puts the system into service [in the EU] under its own name or trademark" 
 - Ggf. wer als Deployer ein System umfunktioniert
- **Weitere Rollen:** Importer, Distributor, Product Manufacturer, EU-Vertreter

Ausführlicher Aufsatz
zum EU AI Act:
vischerlnk.com/3ZkPOYh



Siehe AI Act Check unter
vischerlnk.com/gaira



AI Act: Verbotene Praktiken

• Auswahl

- KI unterschwellig, absichtlich manipulativ oder täuschend eine Person in ihrem Verhalten wesentlich zu beeinflussen (damit sie nicht mehr richtig entscheiden kann) oder um Schwächen von vulnerablen Personen auszunutzen, was zu einem erheblichen Schaden für sie führen kann
 - Gebräuchliche und legitime Praktiken z.B. im Bereich der Werbung, die gesetzeskonform sind, sollen nicht erfasst sein
- KI um aufgrund biometrischer Merkmale Menschen nach ihrer Rasse, ihrer politischen, religiösen oder weltlichen Ansichten, ihrer sexuellen Orientierung oder ihrem Sexualleben einzuteilen
 - Korrelation von Rasse bzw. inneren Werten mit "Äusserlichkeiten"
- Social Scoring oder Profiling mittels KI führt zur nachteiligen Behandlung in Bereichen, die mit den benutzten Daten nichts zu tun haben oder die ungerechtfertigt oder unverhältnismässig ist
 - Zweckgebundene Nutzung von Daten nicht erfasst?
- KI zur Vorhersage ob eine Person straffällig wird, mit Ausnahmen
 - Nicht z.B. Betrugsanalyse von Transaktionen
- Emotionserkennung am Arbeitsplatz und in Bildungseinrichtungen
 - Nicht wo nur zur Sicherheit oder Gesundheit

Beispiel: Chief LOL Officer

Privatkunden Unternehmenskunden Institutionelle Anleger Über uns DE Q

baloise

Versichern Firma gründen Konten, Karten & Finanzierung Anlegen Nachhaltigkeit Kontakt & S...

Der Chief LOL Officer

Laut lachen – gesund arbeiten

BOX DER BALOISE Publiziert 10. Oktober 2024, 08:31

«Chief LOL Officer»: Griesgrämige Angestellte bekommen Memes und Fails

Ist am Arbeitsplatz die Stimmung im Keller, schickt der Versicherungskonzern Baloise jetzt den «Chief LOL Officer» los. Der KI-Bot sendet erheiternde Memes und Videos an mies gelaunte Mitarbeitende.

Quelle: 20 Minuten

The following AI practices shall be **prohibited**:
 ... the placing on the market, the putting into service for this specific purpose, or the **use of AI systems to infer emotions** of a natural person **in the areas of workplace** and education institutions, except where the use of the AI system is intended to be put in place or into the market for medical or safety reasons;

Die für die rechtliche Prüfung zuständigen Stellen wussten offenbar nichts davon ...

AI Act: Hoch-Risiko KI-Systeme

- **Sicherheitskomponenten** von **Produkten** für die es bereits Konformitätsbeurteilungen durch Dritte braucht (gem. Liste)
 - Z.B. Medizinalgeräte, Spielzeug, Funkgeräte, Aufzüge
- Liste von **weiteren Use Cases** (nur teilweise privater Sektor)
 - KI zur biometrischen Emotionserkennung/Kategorisierung → Nicht erfasst ist die biometrische Authentifizierung
 - KI als Sicherheitskomponente bei kritischer Infrastruktur
 - KI für Beurteilungen im Ausbildungsbereich
 - KI zur Beurteilung von Bewerbern und Arbeitnehmern oder diese im Einzelnen betreffende Entscheide (z.B. Arbeitszuteilung)
 - KI zur Steuerung des Zugangs zu wichtigen öffentlichen Diensten und zur Gesundheitsversorgung oder Notfalldiensten → Nicht Erkennung von Emotionen anhand von Texten
 - KI zur Bonitätsbeurteilung und Tarifierung best. Versicherungen → Aber nicht der "Robo-Doc" → Medizinalgerät

AI Act: Wichtigste Pflichten

- Pflichten bei **Hoch-Risiko KI-Systemen** (Auswahl)
 - **Provider:** Risiko- und Qualitätsmanagement, Datenqualität, Konformitätsprüfung, Registrierung, EU-Vertreter, Anleitung, Dokumentation, Incident Monitoring inkl. Meldepflichten
 - **Deployer:** Befolgung der Anleitung, nur geeigneter Input, menschliche Überwachung, Meldepflichten, Transparenzpflichten
- Pflichten bei **anderen KI-Systemen**
 - **Provider:** Hinweis auf Interaktion mit KI, Wasserzeichen
 - **Deployer:** Hinweis auf biometrische Emotionserkennung, Deep Fakes und KI-generierte und automatisch publizierte Inhalte von öffentlichem Interesse müssen als solche erkennbar sein
- Weitere Regeln bestehen für **Allzweck-KI-Modelle**

Ausser Pflicht zur Förderung der "AI Literacy" keine allgemeinen KI-Verhaltenspflichten

Jeden Use Case prüfen, welche Rolle die Organisation hat und ob sie damit im Anwendungsbereich des AI Act liegt

Ist es bald aus mit Daisy?



Die Anbieter stellen sicher, dass KI-Systeme, die für die direkte Interaktion mit natürlichen Personen bestimmt sind, so konzipiert und entwickelt werden, dass die betreffenden natürlichen Personen informiert werden, dass sie mit einem KI-System interagieren, es sei denn, dies ist aus Sicht einer angemessen informierten, aufmerksamen und verständigen natürlichen Person aufgrund der Umstände und des Kontexts der Nutzung offensichtlich.



Ausnahmen nur für
gesetzlich zugelassene
Systeme ...

Quelle: bild.de

AI Act: Anwendungsbeispiele

Fall	Provider	Deployer
Ein Unternehmen in der Schweiz stellt Mitarbeitenden ChatGPT oder Copilot zur Verfügung. Sie verwenden es, um E-Mails, Vorträge, Blogs, Zusammenfassungen und Übersetzungen zu erstellen und Bilder zu generieren. Nutzung auch in der EU.	Nein	Ja
Ein Unternehmen in der Schweiz hat ein unternehmenseigenes Chat-Tool basierend auf einem LLM entwickeln lassen und setzt es intern ein, um E-Mails, Vorträge, Blog-Beiträge, Zusammenfassungen, Übersetzungen und andere Texte zu erstellen und Bilder zu generieren. Es ist geplant, dass auch Personen in der EU die KI-generierten Inhalte erhalten (z.B. als E-Mails oder Texte auf der Website).	(Ja)	Ja
Ein Unternehmen in der Schweiz nutzt ChatGPT oder Copilot, um Unterlagen von Bewerbenden für Stellen in der Schweiz auf etwaige Probleme hin zu analysieren. Die Ergebnisse bleiben intern. Die Bewerbenden kommen auch aus der EU.	Nein	Nein
Ein Unternehmen in der Schweiz stellt auf seiner Website einen selbsterstellten Chatbot zur Beantwortung von allgemeinen Anfragen zum Unternehmen bereit. Die Website (mit dem Chatbot) richtet sich auch an Personen in der EU.	(Ja)	Ja
Ein Unternehmen in der Schweiz nutzt das Produkt bzw. den Service eines Dritten, um den Chatbot auf seiner Website zu realisieren. Diesem werden Inhalte des Unternehmens in der Form einer Datenbank zur Verfügung gestellt (RAG). Das Unternehmen gibt nicht an, von wem der Chatbot stammt.	(Nein)	Ja
Das Unternehmen gibt auf der Website an, von wem der Chatbot stammt ("powered by ...").	Nein	Ja
Ein Unternehmen in der Schweiz setzt lokal ein LLM ein, um Texte zu transkribieren, die auch für die EU bestimmt sind. Das Python-Skript überträgt es von einer kostenlosen Vorlage aus dem Internet unverändert in den eigenen Computer.	Nein	Ja
Ein Unternehmen in der Schweiz setzt einen auch für Kunden in der EU angebotenen Service eines US-Dienstleisters ein, um damit Avatare für Schulungsvideos zu generieren.	Nein	Ja

AI Act: Anwendungsbeispiele

Ein KIS wird benutzt, um Anrufer anhand ihrer Stimme zu identifizieren, damit ihnen Zugang gewährt werden kann bzw. sie bei Anrufen als autorisiert erkannt werden können

Aufgezeichnete Kundengespräche werden von einem KIS protokolliert und ohne Bewertung zusammengefasst

Ein KIS wird benutzt, um Kunden anhand ihres Verhaltens für besondere Marketingaktivitäten auszuwählen

Ein KIS empfiehlt Kunden Anlagen (z.B. im E-Banking)

Ein KIS wertet Kundengespräche im Call Center auf verschiedene Aspekte hin aus, einschliesslich der Zufriedenheit der Kunden, um dem Supervisor Hinweise zu geben, welche Gespräche er für Massnahmen zur Qualitätsverbesserung genauer unter die Lupe nehmen sollte

Die Bonität eines Kunden wird durch ein KIS beurteilt

Ein KIS wird benutzt, um Kundentransaktionen zu identifizieren, bei welchen ein erhöhtes Risiko von Geldwäscherei besteht

Erwartungen der FINMA

1. Es müssen klare **Rollen und Verantwortlichkeiten** sowie **Risikomanagementprozesse** definiert und implementiert werden. **Die Verantwortung für Entscheidungen kann nicht an KI oder Drittparteien delegiert werden.** Alle Beteiligten müssen über genügend **Know-how im Bereich KI verfügen.**
2. Bei der Entwicklung, der Anpassung und in der Anwendung von KI ist sicherzustellen, dass die **Ergebnisse hinreichend genau, robust und zuverlässig sind.** Dabei sind sowohl die Daten als auch die Modelle und die Resultate kritisch zu hinterfragen.
3. Die **Erklärbarkeit der Resultate** einer Anwendung sowie die **Transparenz über deren Einsatz** sind je nach Empfänger, Relevanz und Prozessintegration sicherzustellen.
4. Nicht begründbare **Ungleichbehandlung ist zu vermeiden.**



Worum es der
FINMA dabei geht:

<https://vischerlnk.com/3MRHurk>

Worauf weiter achten?

- Risikobasierter Ansatz: Welche **negativen Auswirkungen** kann ein KI-Einsatz haben? Welche davon sind erheblich?
- Massnahmen für **KI-Governance** (Weisungen, Standards, Dokumentation, AKV) und bankinternes Know-how zum Thema
 - Ferner: Verzeichnis der Anwendungen, Risikobeurteilung, Testen, technische Kontrollen, Monitoring, Reporting
 - Fehlen: Tests, Dokumentation (z.B. Bias, Erklärbarkeit, Fairness)
- **Modell-Risiken** kennen (nicht nur klassische Rechtsthemen), auch bei analytischen Modellen (es gibt nicht nur LLM!)
 - Weitere Risikofelder: Einsatz Dritter, Recht, Reputation, Strategie, IT/Daten/Cyber, Finanzen, Marktverhalten
 - Risikobeurteilung sollte technologieneutral bleiben

Dokumentation des Einsatzes von KI?

- **Anwendungszweck**
- Auswahl und Vorbereitung der **Daten** (Quellen, Qualität, Eignung, Integrität, Richtigkeit, Relevanz, Bias, Einheitlichkeit)
- Auswahl des **Models**
- **Key Performance Indicators**
- **Annahmen**
- **Beschränkungen**
- **Testen** und technische/organisatorische **Schutzmassnahmen**
- **Erklärbarkeit** und Bias
- **Fallback**

Kritische Fragen zur Governance

1. Wissen wir, **wo** wir (relevante) KI im Unternehmen einsetzen?
2. Haben wir Aufgaben, Kompetenzen und Verantwortlichkeiten ("**AKV**") in Sachen KI *lege artis* geregelt, inkl. Ownership?
3. Haben wir die nötige **Fachkompetenz** zum Einsatz von KI?
4. Haben wir unsere Mitarbeitenden im Umgang mit KI **geschult**?
5. Haben wir zu uns passende **Vorgaben** zum Umgang mit KI?
6. Wird KI **kontrolliert** eingeführt und der Einsatz **überwacht**?
7. Betreiben wir ein **Risiko-Management** im Bereich KI, das operationelle, finanzielle, rechtliche und reputative Risiken abdeckt? Wer? Wo sind die grössten Risiken? Massnahmen?
8. Ist die **GL** bzw. **Bankrat** in riskantere Vorhaben eingebunden?

Risikobasierter Ansatz im Bereich KI

- Nicht jede Anwendung von KI hat **relevante** oder gar hohe **Risiken**
 - Fokussieren, Experimente zulassen
- Risiken **strukturiert prüfen** und dokumentieren
- Augenmerk auf **Risiko-Trigger**, z.B.
 - Biometrie, Emotionen, KI betreffend Personal, sensitive Personendaten, Berufs- und Geschäftsgeheimnisse, urheberrechtlich geschützte Inhalte, Einsatz von Service-Providern, KI-Angebote für Dritte, Automatisierung wichtiger Entscheide, Bias, Training

The screenshot displays the GAIKA Score tool interface. At the top, it asks 'Prüfen Sie Ihr GenKI-Projekt auf Risiken' and 'Gibt es massenhafte hohe Risiken für das Unternehmen?'. Below this is a table with columns for 'Anwendungszweck', 'Risikofaktor', 'Anwendungsbereich', 'Anwendung', and 'Risikostufe'. The table contains several rows of data with colored cells (green, yellow, red) indicating risk levels. To the right of the table is a 'Risk Radar' chart showing a red dot in the high-risk area. Below the table are sections for 'Compliance housing questions' and 'Risk questions (25)'. The bottom right corner of the screenshot shows the GAIKA Score logo and the URL 'vischerlnk.com/gaira'.

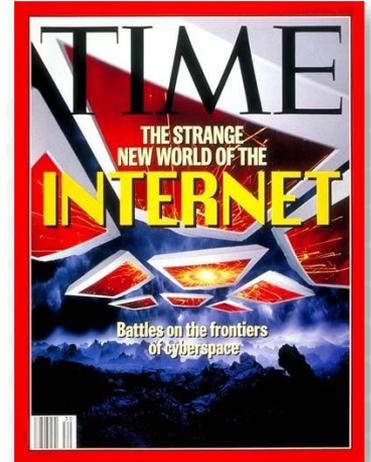
vischerlnk.com/ki-risikocheck

Diskussion zur Governance

- **Definition:** Was ist KI? Was wird als KI erfasst? De-minimis?
- **Inventar:** KI-Infrastruktur? KI-Tools? KI-Modelle? Use Cases?
- **AKV:** Wer prüft KI-Regulierung? Wer lässt Use Cases zu?
- **Weisung:** KI-Weisung ja/nein? Was erlauben? Überwachen?
- **Know-how:** Mitarbeitende? Fachpersonen? Management?
- **Dokumentation, Tests, Monitoring:** Use Cases, Modelle
- **Prüfprozess:** Jede KI? Eigener Prozess? → CCRA, GAIRA
- **Risikomanagement:** Risikoklassen? Methodik? → GAIRA
- **Extern:** KI-Zusatz für Verträge? DSE oder AGB anpassen?

Wie geht es allgemein weiter?

- Die Materie mag heute für viele von uns **unheimlich und ungewohnt** sein
 - Das war das Internet auch mal
- Dies führt zum **Ruf nach mehr Regulierung** und "Ethik"
 - Themen wie Transparenz, Diskriminierung, Erklärbarkeit, Human-in-the-Loop
 - "EU AI Act" als typische Reaktion darauf
 - Schweizer Regulationsbedarf?
- Aber: Wir werden uns **daran gewöhnen**
 - Wir werden unsere bestehenden Gesetze und Prozesse darauf anwenden; es wird so normal werden wie das Internet



25. Juli 1994 (time.com, Titelseite:
James Porto)

VISCHER

Danke für Ihre Aufmerksamkeit!

Fragen: david.rosenthal@vischer.com, lucian.hunger@vischer.com

Zürich

Schützengasse 1
Postfach
8021 Zürich, Schweiz
T +41 58 211 34 00

www.vischer.com

Basel

Aeschenvorstadt 4
Postfach
4010 Basel, Schweiz
T +41 58 211 33 00

Genf

Rue du Cloître 2-4
Postfach
1211 Genf 3, Schweiz
T +41 58 211 35 00

Mehr zum Thema:
vischer.com/ki