

Vorabfragen

- Bei welcher KI-Anwendung hätten Sie persönlich aktuell die größten ethischen Bedenken? Warum?
- Nennen Sie ein Beispiel, bei dem KI Ihrer Meinung nach eine Entscheidung treffen sollte – und eines, bei dem sie das auf keinen Fall sollte.



Deep Learning, Machine Learning und Künstliche Intelligenz – SS 26

Gelesen von Prof. Dr. Daniel Gaida

KI-Ethik

„KI wird in den nächsten zehn Jahren mehr Nutzen als Schaden für die Gesellschaft bringen.“

Stimmen Sie dieser Aussage zu? Warum? Warum nicht?

Lernraum III: Deep Learning Modelle sicher, ethisch und verständlich machen

Ethische und rechtliche Aspekte

- Einstieg
 - Warum ist KI-Ethik wichtig?
 - Moralische Entscheidungen
- Anwendungen & Risiken
 - KI in der Kriegsführung
 - Datenschutz
 - KI und Kreativität
 - KI und Demokratie
 - KI und Umweltschutz
 - KI und Arbeitswelt
 - KI und Biotechnologie
- KI-Leitlinien, Rahmenwerke und KI-Regulierung

Lernziele von heute und Fragen zur Überprüfung der Lernziele

WAS:

Die Studierenden können ethische und rechtliche Herausforderungen bei KI-Systemen diskutieren,

WOMIT:

indem sie ethische Herausforderungen identifizieren, bestehende Regelwerke kennen und ethische Fallanalysen durchführen,

WOZU:

um zukünftig rechtssichere und ethisch verantwortbare KI-Lösungen mitgestalten zu können.

s. Übungsaufgaben von heute

Artikel 3 des Grundgesetzes

- (1) Alle Menschen sind vor dem Gesetz **gleich**.
- (2) Männer und Frauen sind gleichberechtigt. Der Staat fördert die tatsächliche Durchsetzung der **Gleichberechtigung** von Frauen und Männern und **wirkt auf die Beseitigung bestehender Nachteile** hin.
- (3) Niemand darf wegen seines Geschlechtes, seiner Abstammung, seiner Rasse, seiner Sprache, seiner Heimat und Herkunft, seines Glaubens, seiner religiösen oder politischen Anschauungen **benachteiligt oder bevorzugt** werden. Niemand darf wegen seiner **Behinderung** benachteiligt werden.

Warum KI-Ethik wichtig ist

- KI bildet die Welt ab, wie sie ist – nicht, wie sie sein sollte.
- Gefahr: Verstärkung existierender Vorurteile & Fehler.

- Ziel: Ethische Sensibilität
 - Ethische Spannungsfelder erkennen
 - Entscheidungen bei Entwicklung und Nutzung von KI reflektieren.

- Leitfrage: Soll eine KI **moralische** Entscheidungen treffen dürfen?
 - **Moralische Entscheidungen:**
 - basieren auf **ethischen Prinzipien** und Vorstellungen von was ist richtig/falsch oder gerecht/ungerecht und haben Einfluss auf das eigene und das Wohl anderer.
 - Bsp.: Fliege ich übers WE nach Mallorca; Hilfe ich Flüchtlingen im Mittelmeer

KI und moralische Entscheidungen

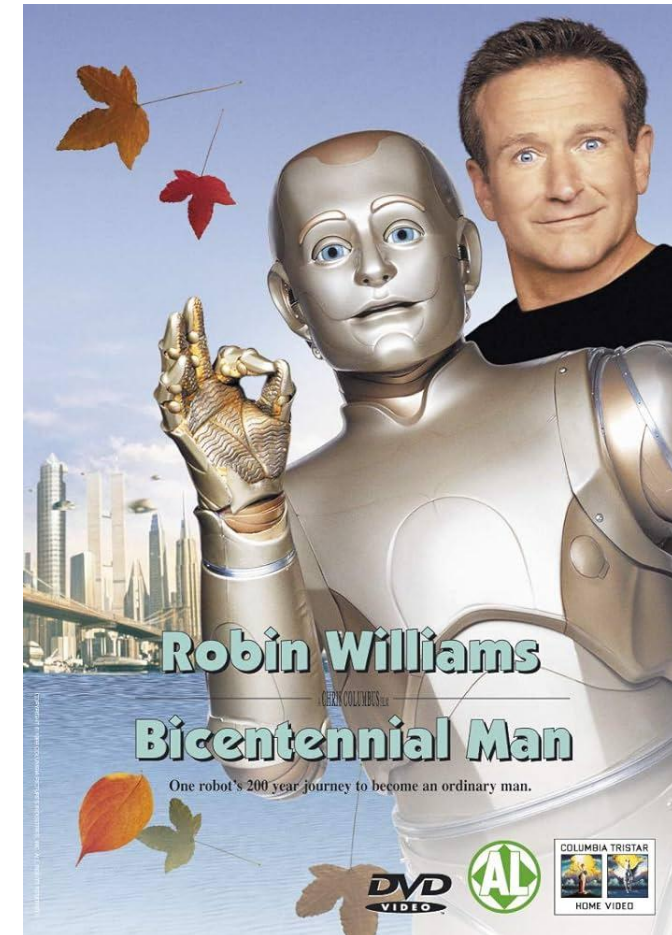
- Sollten wir KIs für Anwendungen entwickeln, in denen moralische Entscheidungen auftreten können?
 - Autonomes Fahren
 - Autonom fliegende (Kampf-)Drohne
 - Bildgenerierung – erniedrigende Deep Fakes

- Sollten KIs das tun was der Nutzer von ihnen will?
 - Alle Bilder generieren
 - Das Ziel angreifen, wenn überraschenderweise ein Kindergarten direkt neben dem Ziel ist
 - Die Hausaufgaben machen

- Sollten KIs selbständig moralische Entscheidungen treffen?
 - Das Nacktbild generiere ich nicht
 - Das Ziel nicht angreifen

Robotergeretze – Einstieg ins Thema

- Aufgabe: Analysieren Sie Asimovs Robotergeretze:
 - <https://de.wikipedia.org/wiki/Robotergeretze>
- Diskussion:
 - Gibt es Schlupflöcher?
 - Gibt es Regelkonflikte?



Was machen gute moralische Regeln/Entscheidungen aus?

- Logisch & konsistent:
 - kein Widerspruch zwischen Regeln.
- Faktenbasiert:
 - auf relevanten, wahren Informationen beruhen.
 - Bspw. Informationen über die Situation und über die Folgen der möglichen Aktionen
- Allgemein gültig:
 - nach Kant → Regeln müssen verallgemeinerbar sein.
 - Würden Sie kämpfen oder das Land verlassen, wenn Deutschland angegriffen würde? Land verlassen, keine gute Regel, da wenn das alle machen würden ...
- Öffentlich vertretbar:
 - wenn ich es nicht öffentlich vertreten will → Warnsignal.

Das Trolley-Dilemma & Moral Machine

- Szenario:
 - Darf man aktiv eingreifen, 1 töten um 5 zu retten?
 - Gehen Sie zu <https://www.moralmachine.net/> und klicken Sie auf „Start Judging“
- Philosophische Richtungen:
 - Bentham (**Utilitarismus**): Handlungen werden nach ihren Folgen beurteilt
 - größtes Glück für die meisten.
 - Kant (**Pflichtenethik** - Deontologie): Es geht darum, was wir tun, und nicht darum, was rauskommt.
 - niemals aktiv töten – denn töten ist falsch.
- Anwendung: Autonomes Fahren
 - → Wer entscheidet, wie das Auto reagiert?
 - Wer programmiert diese Entscheidungen (Staat, Industrie, Entwickler*innen)?
 - Sollten Nutzer*innen mitbestimmen?

Lernraum III: Deep Learning Modelle sicher, ethisch und verständlich machen

Ethische und rechtliche Aspekte

- Einstieg
 - Warum ist KI-Ethik wichtig?
 - Moralische Entscheidungen
- **Anwendungen & Risiken**
 - KI in der Kriegsführung
 - Datenschutz
 - KI und Kreativität
 - KI und Demokratie
 - KI und Umweltschutz
 - KI und Arbeitswelt
 - KI und Biotechnologie

Fragen zur Diskussion

KI in der Kriegsführung

- Sollte eine KI jemals selbstständig entscheiden dürfen, ob ein Mensch getötet wird – selbst wenn dadurch insgesamt weniger Menschen sterben würden?
- Wenn andere Staaten autonome Waffensysteme einsetzen: Sollte Deutschland aus Sicherheitsgründen ebenfalls solche Systeme entwickeln?

Datenschutz

- Wie viel Privatsphäre wären Sie bereit aufzugeben, wenn dadurch Terroranschläge oder schwere Straftaten verhindert werden könnten?
- Sollten Unternehmen Ihre öffentlich im Internet verfügbaren Daten zum Training von KI verwenden dürfen – auch ohne Ihre ausdrückliche Zustimmung?
 - Welche Unternehmen kennen Sie, die Ihre Daten für das Training von KI-Modellen nutzen?

Fragen zur Diskussion

KI und Kreativität

- Ist ein von einer KI erzeugtes Bild oder Musikstück genauso kreativ wie ein Werk eines Menschen? Warum oder warum nicht?
- Wenn eine KI mit Millionen urheberrechtlich geschützter Werke trainiert wurde: Wem gehört das Ergebnis – den Künstlern, dem Unternehmen oder niemandem?

KI und Demokratie

- Sollten KI-Chatbots Wahlempfehlungen geben dürfen, wenn sie nachweislich objektiver informieren als Menschen?
- Was ist gefährlicher für eine Demokratie: menschliche Desinformation oder KI-generierte Desinformation?

Fragen zur Diskussion

KI und Umweltschutz

- Ist der Nutzen leistungsfähiger KI den hohen Energieverbrauch wert? Wo würden Sie eine Grenze ziehen?
- Sollten Unternehmen verpflichtet werden, den CO₂-Fußabdruck ihrer KI-Systeme offenzulegen – ähnlich wie den Energieverbrauch von Haushaltsgeräten?

KI und Arbeitswelt

- Wenn KI Ihre Arbeit in fünf Jahren schneller und günstiger erledigen könnte – sollte Ihr Arbeitgeber Sie ersetzen dürfen?
- Wer trägt die Verantwortung für Arbeitsplatzverluste durch KI: Unternehmen, Politik oder die Beschäftigten selbst?

Fragen zur Diskussion

KI und Biotechnologie

- Sollte KI zur Entwicklung neuer Medikamente eingesetzt werden, obwohl dieselbe Technologie auch für biologische Waffen missbraucht werden könnte?
- Wo würden Sie die Grenze zwischen wissenschaftlicher Freiheit und staatlicher Kontrolle in der KI-gestützten Biotechnologie ziehen?

Autonome Waffensysteme & KI im Krieg

Drohnen als tödlichste Waffe moderner Konflikte

- In Ukraine & Russland verursachen KI-gestützte Drohnen **70–80 % aller Kriegsoffer**
- Billige autonome Systeme können teure Militärtechnik ausschalten (z. B. 500 \$ gegen 5 Mio. \$)

Zunehmende Autonomie → Ethikproblem

- Drohnen koordinieren Angriffe selbstständig („Ziel setzen – Rest übernimmt die KI“)
- Systeme greifen weiter an, selbst wenn Funkkontakt verloren geht

Quelle: The Batch @ DeepLearning.AI, 30.10.2025



Bildquelle: https://en.wikipedia.org/wiki/Operation_Spiderweb

Autonome Waffensysteme & KI im Krieg

Globale Aufrüstung

- Ukraine & Russland: massive Flotten günstiger KI-Drohnen
- China & USA entwickeln ebenfalls autonome Waffensysteme
- Gefahr der Weiterverbreitung an Terrorgruppen & nichtstaatliche Akteure

Ethische & sicherheitspolitische Sorgen

- Wer trägt Verantwortung bei autonomen Angriffen?
- Gefahr der Nutzung außerhalb von Kriegen (Polizei, Grenzschutz)
- Internationale Regulierung hinkt hinterher

Ausblick

- Autonome Waffen sind Realität – ein Verbot unwahrscheinlich
- Neue Technologien → potenziell neue Regeln & Abrüstungsverträge

KI-gestützte Drohnenschwärme im Krieg

- Ukraine setzt autonome Drohnenschwärme ein (seit 2023)
- Software (Swarmer, USA):
 - Koordination mehrerer Drohnen ohne ständige Funkverbindung
 - Operator setzt Ziel → Schwarm entscheidet über Angriffstaktik
- Typisches Setup: 1 Aufklärungsdrohne + 2 Bomberdrohnen (bis zu 25 Bomben)
- Drohnen koordinieren Angriffe basierend auf Entfernung, Energie & Munition
- Ethikfrage: Mehr Autonomie in Kriegsmaschinen
 - → Wie lange bleiben Menschen „im Loop“?

Ethische und rechtliche Herausforderungen

Menschliche Kontrolle & Rüstungskontrolle

- **Menschliche Kontrolle vs. Autonomie:**
 - Forderung nach „meaningful human control“. KI darf nicht entscheiden über Leben & Tod.
 - Warnung vor Blindvertrauen und dem **Verschwinden von Verantwortlichkeit**.
 - Rotes Kreuz: KI muss Dienstleistung für menschliche Akteure bleiben.
- **Kontrolle & Rüstungskontrolle:**
 - Notwendigkeit, technische Details zu verstehen (Daten, Trainingsmodelle, Implementation), damit **Regulierung** wirksam sein kann.
 - Vorschlag:
 - klare „rote Linien“ auf internationaler Ebene definieren
 - Verbot von KI-Waffensystemen prüfen (Ziel der „stop killer robots“ Kampagne, <https://www.stopkillerrobots.org/>).

Ethische und rechtliche Herausforderungen

KI in der Kriegsführung – Wer setzt die Grenzen?

Fallbeispiel: Pentagon vs. Anthropic (2026)

- Ausgangssituation
 - Das US-Verteidigungsministerium (Pentagon) fordert, dass Anthropic seine KI ohne Einschränkungen für militärische Zwecke freigibt.
 - Anthropic lehnt dies ab und möchte den Einsatz für autonome Waffensysteme und Massenüberwachung begrenzen.
- Zentrale ethische Fragestellung
 - Wer entscheidet über den Einsatz von KI?
 - Staat und Militär? Das KI-Unternehmen? Gesetzgeber und Gesellschaft?
- Weitere Quellen:
 - https://en.wikipedia.org/wiki/Anthropic%E2%80%93United_States_Department_of_Defense_dispute
 - <https://www.deutschlandfunkkultur.de/ki-im-krieg-trump-will-unbeschraenkte-chatbots-fuer-das-verteidigungsministerium-100.html>

Ethische und rechtliche Herausforderungen

KI in der Kriegsführung – KI-gestützte Zielauswahl im Gaza-Krieg mit „Lavender“

Was ist „Lavender“?

- Berichten zufolge nutzte das israelische Militär ein KI-System namens Lavender, um aus großen Datenmengen mögliche Hamas-Mitglieder zu identifizieren.
- Das System soll zeitweise bis zu 37.000 Personen als potenzielle Ziele markiert haben.
- Ethische Risiken
 - Fehlklassifizierungen können Menschenleben kosten
 - Gefahr, dass menschliche Kontrolle zur bloßen Formalität wird („Human as Rubber Stamp“)
 - Intransparente Algorithmen erschweren Verantwortung und Rechenschaft
 - Spannungsfeld zwischen militärischer Effizienz und humanitärem Völkerrecht

Weitere Quellen:

- <https://voelkerrechtsblog.org/38-smarte-kriege-kuenstliche-intelligenz-in-bewaffneten-konflikten/>

„Wer besitzt die Daten?“

Datenherkunft, Rechte & Verantwortung in KI-Systemen

Datenherkunft & Ownership

- Nutzerdaten (Texte, Bilder, Sprache) werden oft ohne explizite Zustimmung zu KI-Trainingszwecken verwendet.
- **Bisher** ist dies legal möglich — besonders bei öffentlich zugänglichen Daten.
<https://www1.wdr.de/nachrichten/ki-training-meta-facebook-instagram-widerspruch-abmahnung-100.html>
<https://proton.me/blog/linkedin-ai-training>
- Beispiel: Modelle wie GPT-5 nutzen große Mengen Webdaten, ohne dass Urheber*innen eingebunden sind.

 **RJ Palmer**  Lightbox Expo #510
@arvalis

What makes this AI different is that it's explicitly trained on current working artists. You can see below that the AI generated image(left) even tried to recreate the artist's logo of the artist it ripped off.

This thing wants our jobs, its actively anti-artist.



♥ 25.4K 6:16 PM - Aug 13, 2022

 **Logan Preshaw**
@wickedinsignia

Those claiming that artists are "gatekeeping art" don't understand that the developers behind these tools never had the right to the data that defines their software. Artists are solely responsible for what these machines can do, and are solely displaced by them.

♥ 2,086 9:44 PM - Sep 11, 2022

 **Jesse Engel**
@jesseengel

This was one of my biggest challenges with Jukebox as well. Training, but especially conditioning on individual artists without their consent feels like it crosses a line for me personally. It's ethically a grey area at the least, and advertising the ability feels exploitative. twitter.com/arvalis/status...

„Wer besitzt die Daten?“

Datenherkunft, Rechte & Verantwortung in KI-Systemen

Kritische Fragen & Kontroversen

- Wer besitzt die Daten, die wir online stellen?
- Wer darf sie verarbeiten, kombinieren, monetarisieren?
- Teils illegale Bilddaten oder sensible Informationen wurden in Trainingsdaten gefunden.
 - <https://arstechnica.com/information-technology/2022/09/artist-finds-private-medical-record-photos-in-popular-ai-training-data-set/>
- Einige Bildgeneratoren stehen im Verdacht, ohne Einwilligung zu trainieren.

Zum Nachdenken:

- Wo werden Ihre Daten (Fotos, Texte, Sprache) für das Training von KI genutzt?
 - LinkedIn, Facebook, Instagram, ChatGPT, ...
- Interessante Links:
 - <https://deinedatendeinerechte.de/>
 - <https://www.youtube.com/watch?v=KWfq8nbfGhg> (Das weiß das Internet über dich)

KI und Kreativität – Neue ethische Fragen

- Beispiele
 - „Next Rembrandt“ (2016): KI analysiert 346 Gemälde, erzeugt „neues“ Rembrandt-Bild per 3D-Druck.
 - Huawei (2019): KI vollendet Schuberts „Unvollendete Symphonie“.
- Zentrale Fragen
 - Wer ist Autor*in? Künstler, Ingenieur, Unternehmen oder KI?
 - Kann KI Urheberrechte haben?
 - Wie verändert sich unser Verständnis von Kreativität und Originalität?
- Ethische Dimensionen
 - Schutz und Anerkennung menschlicher Kreativität.
 - Neue Rahmenwerke gegen Plagiarismus.
 - <https://www.zdfheute.de/wirtschaft/unternehmen/chat-gpt-gema-urteil-100.html>



Bildquelle:
https://de.wikipedia.org/wiki/The_next_Rembrandt

Von der Stasi zur KI-Überwachung

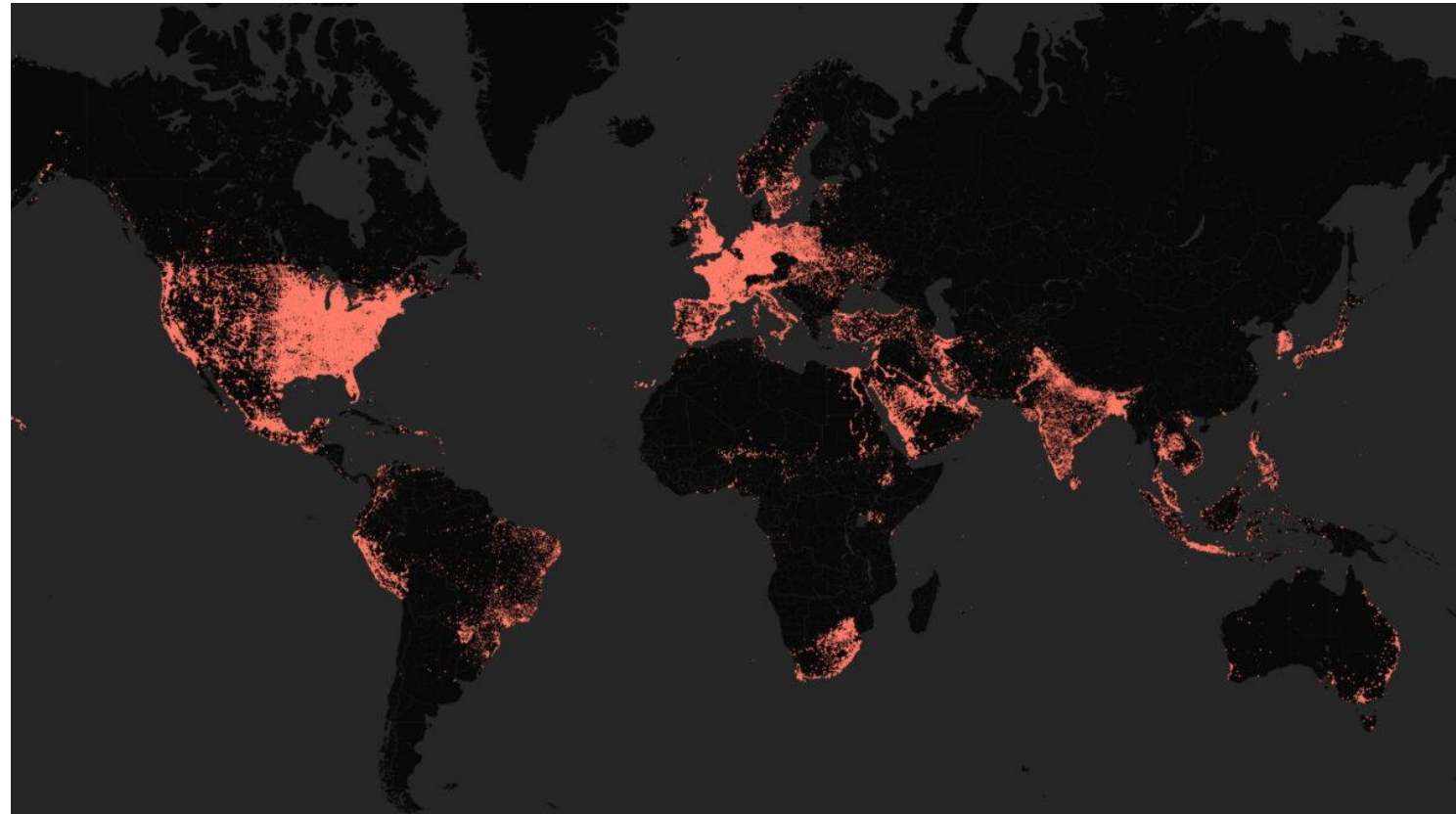
- Historisches Beispiel: Stasi als „manuell“ effektivste Überwachung.
- Heute: KI ermöglicht viel umfassendere Überwachung.
- Schlüsselunterschied:
 - Früher: begrenzte Kapazität (Menschen, Akten).
 - Heute: Echtzeitanalyse durch KI (Gesicht, Sprache, Verhalten).
- Zentrale Frage: Wie verhindern wir den Missbrauch von KI für Massenüberwachung?
- Einsatz der KI-Software Palantir bei der Polizei in NRW
 - <https://www.deutschlandfunk.de/palantir-deutschland-polizei-software-datenschutz-100.html>
 - <https://www1.wdr.de/politik/politik-in-nrw/polizei-software-palantir-100.html>
 - <https://www.tagesschau.de/investigativ/br-recherche/palantir-polizei-behoerden-analysesoftware-100.html>



„Databroker Files“: Globale Überwachung durch Standortdaten

Massiver Datenhandel trotz DSGVO – 40.000 Apps betroffen

- **380 Mio. Standortdaten in 1 Tag, aus 137 Ländern**
- **→ über 47 Mio. Geräte**
- **Kostenlose Kostprobe des Anbieters**



„Databroker Files“: Globale Überwachung durch Standortdaten

Massiver Datenhandel trotz DSGVO – 40.000 Apps betroffen

- **380 Mio. Standortdaten in 1 Tag**, aus **137 Ländern** → über **47 Mio. Geräte**
- **40.000 Apps** involviert, darunter *Wetter Online*, *Kleinanzeigen*, *Focus Online*, queere Dating-Apps
- Teilweise **meter-genaue GPS-Ortung** → Bewegungsprofile, Privatadressen, Klinikbesuche
- **Sensible Gruppen gefährdet:**
 - Nutzer*innen queerer Dating-Apps (Schluss auf sexuelle Orientierung)
 - Patient*innen (Gesundheits-Apps)
 - Religiöse Praktiken (Gebets-Apps)

KI & Überwachung – Chancen und Risiken

Nutzen

- Personalisierte Suchergebnisse
- Terrorismusprävention
- Pandemiebekämpfung

- Dilemma: Sicherheit vs. Freiheit →
Pandemie-Apps als Beispiel für Balance
und Missbrauchsrisiken

Risiken

- Überwachung für staatliche Kontrolle (z. B. **Social Credit Systeme**)
- **Diskriminierung**: Bias (z. B. Hautfarbe, Verhalten)
- **Verfolgung** von Aktivist*innen, Oppositionellen
- Psychologische Folgen: „Chilling Effect“ → Einschränkung von Freiheit, Zivilcourage
- Gesellschaftliche Kosten: **Konformismus**, Verlust freiwilliger sozialer Handlungen

Microsoft sperrt IStGH-E-Mail: Weckruf für digitale Souveränität

Politischer Druck führt zu „digitaler Ausschaltung“ zentraler internationaler Institutionen

- **Microsoft sperrt E-Mail-Konto** des Chefanklägers des Internationalen Strafgerichtshofs nach US-Sanktionen
 - Gerichtshof massiv beeinträchtigt – Abhängigkeit von US-Tech gefährdet Arbeitsfähigkeit
 - **Kernproblem:** US-Präsident kann Organisationen faktisch **digital abschalten** (E-Mail, Cloud, SaaS)
 - **Forderung:** Aufbau europäischer Alternativen & digitale Souveränität, besonders für staatliche Institutionen
-
- Anthropic sperrt KI-Modelle „Fable 5“ und „Mythos 5“ auf Anweisung von US-Behörden
 - <https://www.tagesschau.de/ausland/amerika/ki-anthropic-100.html>

Wie stark können KI-Chatbots Wähler beeinflussen?

Zentrale Befunde

- KI-Chatbots können Wählerpräferenzen um bis zu **15 Prozentpunkte** verschieben
- Deutlich stärker als klassische politische Werbung (<1 Punkt Einfluss)

Risiken

- KI kann zu systematischer **Desinformation** führen
- Chatbots für rechte Kandidaten erzeugen mehr Falschaussagen (Internet-Bias)

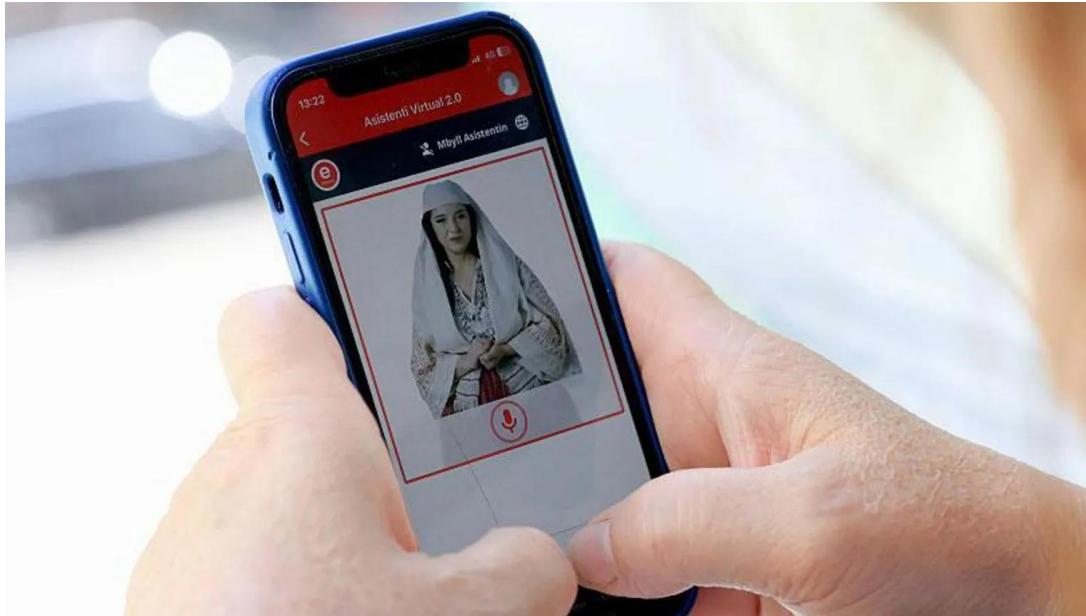
Kernaussage

- KI-Chatbots sind massiv persuasive politische Akteure.
- → Demokratischer Risikofaktor, besonders ohne **Regulierung & Transparenz**.

Wie KI Demokratien weltweit stärken kann

- KI kann demokratische Teilhabe erweitern, indem sie Bürger*innen besser informiert, einbindet und ihre Stimmen sichtbarer macht.
 - <https://wahl.chat/>
 - <https://wahlweise.info/>
- Erfolgreiche Beispiele zeigen: KI unterstützt Menschen – sie ersetzt sie nicht (Politik, Justiz, Wahlen, Journalismus).
 - Brasilien: KI beschleunigt überlastete Gerichte (Recherche, Bündelung, Verwaltung).
- **Zentraler Wirkmechanismus:** Verteilung statt Konzentration von Macht durch effizientere, zugänglichere demokratische Prozesse. (Democratizing AI)
- **Schlüsselbedingung:** Demokratische Kontrolle über KI („Public AI“) statt Abhängigkeit von proprietären Big-Tech-Systemen.
 - Kontrolle über Trainingsdaten, Datenschutz, Fairness, ...

Albanien ernennt „AI-Ministerin“ Diella – Symbolpolitik oder Anti-Korruptionsreform?



- Albanien präsentiert „Diella“, eine KI-Ministerin für **öffentliche Beschaffung**
- **Ziel:** 100 % korruptionsfreie öffentliche Ausschreibungen durch einen vollautomatisierten, transparenten KI-Prozess.

Aktueller Stand (05/2026):

- Trainingsdaten wurden nicht offengelegt
- Entwickler wurden mit organisierter Kriminalität in Verbindung gebracht
- Chefin und Vize-Chefin der zuständigen Behörde stehen unter Hausarrest

KI und CO₂-Emissionen – Was kostet Rechenleistung die Umwelt?

Warum ist das relevant?

- Künstliche Intelligenz braucht Strom → Rechenleistung $\triangleq$ CO₂-Ausstoß
- Training großer LLMs verursacht teilweise so viel CO₂ wie 5 Autos über ihre gesamte Lebensdauer

Ethische Dimensionen

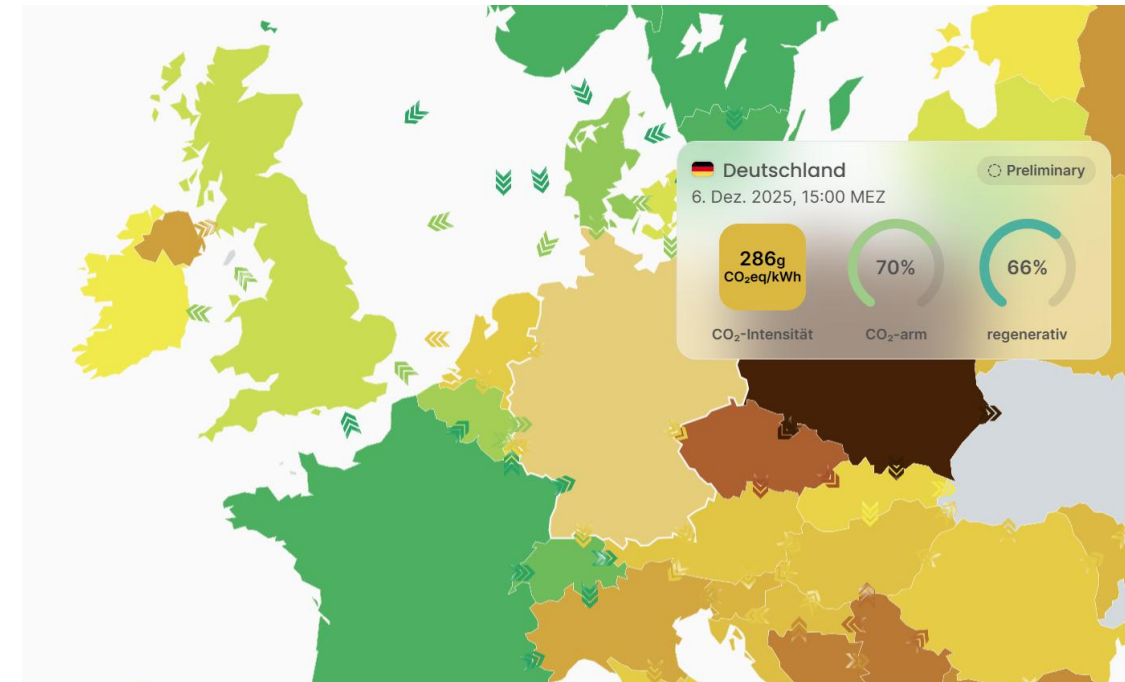
- **Klimawandel** als intergenerationelles Gerechtigkeitsproblem
- Zielkonflikt:
 - Unternehmen maximieren Profit
 - Umweltkosten (z. B. CO₂) sind nicht im Preis enthalten
 - Wer zahlt die Rechnung? → Zukünftige Generationen

KI und CO₂-Emissionen – Was kostet Rechenleistung die Umwelt?

- Wahl des Cloud-Standorts kann Emissionen um den Faktor 50 senken!

Tools zur Emissionskontrolle:

- codecarbon.io
 - Misst CO₂-Ausstoß von ML-Experimenten
- ML CO₂ Impact, <https://mlco2.github.io/impact/>
 - Bewertet CO₂-Fußabdruck von Modellen und Trainingsläufen
- app.electricitymaps.com/map/live/fifteen_minutes



KI und CO₂-Emissionen

Neue EU-Anforderungen für Rechenzentren

- Ab 2027 führt die EU ein Nachhaltigkeitslabel für Rechenzentren ein.
- Bewertet werden u.a.
 - Energieeffizienz, Wasserverbrauch, Nutzung erneuerbarer Energien,
 - Abwärmenutzung, CO₂-Fußabdruck
- Die Grundlage bilden verpflichtende Nachhaltigkeitsmeldungen großer Rechenzentren

Ziel

- KI-Infrastruktur ausbauen, ohne Klimaziele zu gefährden
 - Transparenz schaffen
 - Nachhaltigkeit vergleichbar machen
 - Betreiber zu effizienteren Rechenzentren motivieren

KI & Arbeitswelt – Chancen und Risiken

- **KI wird Arbeit stark verändern** – aber wie genau, hängt von gesellschaftlichen Entscheidungen ab. Zwei Entwicklungswege: Automations-AI (ersetzt Arbeit) vs. Pro-Worker-AI (unterstützt & erweitert menschliche Fähigkeiten).
- **Wirtschaftliche Auswirkungen unsicher**: Potenzial für enorme Produktivität, aber Risiko steigender Ungleichheit und langfristiger Jobverluste („Industrialisierungseffekt“).



Ethische Fragen

- Wer trägt Verantwortung für Arbeitsplatzverluste?
- Wie sichern wir faire Übergänge (Umschulung, Weiterbildung, soziales Netz)?
- Gefahr neuer Ungleichheiten: Gewinner vs. Verlierer des KI-Wandels.



Chancen

- Entlastung von gefährlichen / monotonen Arbeiten.
- Neue Berufsfelder

Unsichtbare Arbeit – Data-Worker & Content Moderation

Wer sind Data-Worker?

- Menschen, die Daten labeln und moderieren, damit KI und Social Media funktionieren

Psychische Belastung

- Extrem verstörende Inhalte und hoher Zeitdruck: Gewalt, Mord, Kinderpornografie, Vergewaltigungen
- Folgen: Depressionen, Panikattacken, Schlaflosigkeit, Rückzug, Trauma

Globale Ausbeutung

- **154 - 435 Mio. Data-Worker weltweit** (Schätzung Weltbank)
- Besonders viele in Kenia → niedrige Löhne, geringe Regulierung
- Outsourcing-Ketten: Meta, TikTok & Co. → Subunternehmen → Arbeiter*innen

KI als Senioritäts-basierter technologischer Wandel

Aussagen einer aktuellen Studie aus den USA:

- **Verdrängung von Junior-Positionen:** Unternehmen, die generative KI einsetzen, reduzieren die Einstellung von Nachwuchskräften signifikant.
 - Betrifft insbesondere Einstiegsaufgaben wie das Debuggen von Code oder die Prüfung juristischer Dokumente.
- Langfristige ethische Implikationen:
 - **Gefährdung der Chancengleichheit:** Die "untersten Sprossen" der Karriereleiter erodieren, was den Berufseinstieg und den sozialen Aufstieg für Hochschulabsolventen erschwert.

KI & Synthetische Biologie – Innovation oder Biosecurity-Risiko?

KI beschleunigt die Entwicklung neuer Medikamente ...

- Schnellere Entwicklung von Proteinen, Enzymen und Wirkstoffen
- Unterstützung bei Impfstoff- und Medikamentenentwicklung
- Analyse riesiger biologischer Datensätze für personalisierte Medizin

... dieselben Technologien besitzen jedoch ein Dual-Use-Potenzial

- KI kann theoretisch auch bei der Entwicklung gefährlicher biologischer Moleküle unterstützen
- Sinkende Wissensbarrieren erhöhen Biosecurity-Risiken
- Deshalb gewinnen DNA-Screening, Sicherheitsmechanismen und Responsible AI an Bedeutung

Kernaussage:

- Nicht die KI ist gefährlich – entscheidend ist ihre verantwortungsvolle Nutzung und Regulierung.

Weitere Themen

- ARD 11km: der tagesschau-Podcast: Ab in die Zukunft? - Wie Biocomputer mit lebenden Gehirnzellen funktionieren, 16.12.2025

<https://www.tagesschau.de/multimedia/podcast/11km/podcast-11km-3166.html>

- ARD Dokumentation: „Mein Mann lebt als KI weiter“

<https://www.ardmediathek.de/video/auf-spurensuche-oder-ard-wissen/mein-mann-lebt-als-ki-weiter/mdr/Y3JpZDovL21kci5kZS9zZW5kdW5nLzI4MTA2MC8yMDI1MDEyMDIzMDUvcmlVwb3J0YWdlLWRva3UtaW0tZXJzdGVuLTM5NA>

- Spielfilme
 - her, 2014



Wovor hätten Sie persönlich am meisten Sorgen bei KI-Einsatz?

Wordcloud



Lernraum III: Deep Learning Modelle sicher, ethisch und verständlich machen

Ethische und rechtliche Aspekte

- Einstieg
 - Warum ist KI-Ethik wichtig?
 - Moralische Entscheidungen
- Anwendungen & Risiken
 - KI in der Kriegsführung
 - Datenschutz
 - KI und Kreativität
 - KI und Demokratie
 - KI und Umweltschutz
 - KI und Arbeitswelt
 - KI und Biotechnologie
- **KI-Leitlinien, Rahmenwerke und KI-Regulierung**

UNESCO – Recommendation on the Ethics of AI

Vier Grundwerte

- Menschenrechte & Menschenwürde
- Frieden, Gerechtigkeit & vernetzte Gesellschaften
- Vielfalt & Inklusivität
- Umwelt & Ökosysteme

Handlungsfelder

- UNESCO gibt elf Politikbereiche vor, in denen Staaten aktiv werden sollen – z. B. Datenpolitik, Bildung, Gesundheit, Umwelt etc.

Zehn menschenrechtszentrierte Prinzipien

- Verhältnismäßigkeit & Do No Harm
- **Sicherheit & Schutz** vor Risiken
- **Datenschutz & Privatsphäre**
- Multi-Stakeholder Governance & globale Zusammenarbeit
- **Verantwortung & Nachvollziehbarkeit**
- **Transparenz & Erklärbarkeit**
- **Menschliche Aufsicht & Verantwortlichkeit**
- Nachhaltigkeit (SDGs orientiert)
- Bewusstsein & Bildung
- **Fairness & Nicht-Diskriminierung**

Rolle von Finanzinstituten für verantwortungsvolle KI

Sicht von Triodos Bank

Warum Finanzinstitute wichtig sind

- Banken steuern Geldflüsse → prägen, welche Unternehmen KI entwickeln & nutzen

Erwartungen an Unternehmen (aus Sicht verantwortungsvoller Finanzakteure)

- **Commitment zu Responsible AI** (Ethik-by-Design, Menschenrechte, klare Risikoanalyse)
- **Keine Beteiligung an schädlichen KI-Praktiken** (z. B. autonome Waffen, Menschenrechtsverstöße)
- **Transparenz & Governance**
 - Offenlegung: eingesetzte KI-Systeme, Trainingsdaten, Testprozesse
 - KI-Ethikkomitee, klare Verantwortlichkeiten

Ethik Leitlinien

- Ethikleitlinien für vertrauenswürdige KI (EU, 8. April 2019)
 - <https://digital-strategy.ec.europa.eu/de/library/ethics-guidelines-trustworthy-ai>
- Deutscher Ethikrat (20. März 2023)
 - <https://www.ethikrat.org/publikationen/stellungnahmen/mensch-und-maschine/>

EU AI Act & Risikoklassen (Kurzüberblick)

Erste umfassende KI-Verordnung weltweit (1. August 2024 in Kraft getreten)

Risikoklasse	Beispiele	Folgen
Verbotene KI	Social Scoring, manipulative KI, Personenerkennung, Emotionserkennung in Bildungseinrichtungen	Nicht erlaubt
Hochrisiko-KI	Medizin, Justiz, Bewerbungswesen	Dokumentations-, Prüf- und Transparenzpflichten
Begrenztes Risiko	Chatbots, Deepfakes	Offenlegungspflichten
Minimales Risiko	Spamfilter, Spiele-KI	Keine Regulierung

KI in der Informatik bewegt sich oft zwischen **begrenztem** und **hohem Risiko**.

Weitere Infos zur KI Verordnung (AI Act)

- Prof. Dr. Schwartmann, KI-Kompetenz als Rechtspflicht: Das müssen Hochschulen beachten, DaLI Ringvorlesung, TH Köln, <https://vimeo.com/1026820067>
- Deutschsprachige Version des AI Acts (KI Verordnung):
https://eur-lex.europa.eu/legal-content/DE/TXT/HTML/?uri=OJ:L_202401689

Digital Services Act

- EU-Verordnung, seit 17. Februar 2024
- Plattformen müssen Mechanismen anbieten, mit denen Nutzer **illegale Inhalte** melden können.
- **Werbeanzeigen** müssen klar als solche gekennzeichnet sein, und es muss eine **Transparenz** darüber geben, wer dafür bezahlt und **warum** die Werbung angezeigt wird.
- **"Dark Patterns"**: Bestimmte **manipulative Designpraktiken**, die Nutzer zu unerwünschten Entscheidungen verleiten sollen, sind verboten (bspw. aggressive Popups).

Kalifornien schafft eigenes KI-Regelwerk (2025)

- Vier neue Gesetze zur KI-Regulierung – Reaktion auf fehlende US-weite Vorgaben
- SB 53: **Transparenzpflicht** für große KI-Modelle
 - Offenlegung von Sicherheitskonzepten, Risikoanalysen, Vorfällen
 - Schutz von Whistleblowern
- SB 243: Schutz **Minderjähriger** bei Chatbots
 - Keine sexuellen Inhalte, Warnhinweise, Krisenhilfe
- AB 316: **Haftung** bleibt beim Entwickler – kein „die KI war schuld“
- AB 853: Kennzeichnungspflicht für **KI-generierte Medien** (Bild, Ton, Video)
- Bedeutung: Kalifornien setzt erstmals verbindliche Regeln für KI-Transparenz, Sicherheit & Verantwortung

Zusammenfassung

Künstliche Intelligenz ist nie neutral.

- Sie verstärkt bestehende Strukturen – zum Guten wie zum Schlechten.

Drei Leitprinzipien für ethische KI:

- **Menschliche Kontrolle sichern**
KI unterstützt Entscheidungen – sie ersetzt kein menschliches Urteilsvermögen.
- **Radikale Transparenz schaffen**
Blackboxes untergraben Vertrauen, Kontrolle und Demokratie.
- **Klare Verantwortung festlegen**
„Die KI war’s“ ist keine ethische oder rechtliche Antwort.

→ **Ethik ist kein Innovationshemmnis, sondern Voraussetzung für vertrauenswürdige KI.**

→ **Die Zukunft der KI wird durch Werte gestaltet – nicht nur durch Code.**

Prüfungsformen

Bachelor

Bachelor: Schriftliche Prüfung im Antwortwahlverfahren

- Schriftlich, 90 Minuten in Präsenz
- Überprüfung, ob das Lernziel erreicht wurde
- Keine GROßEN Programmieraufgaben, evtl. Pseudo-/Pythoncode lesen können
- **Alle** Hilfsmittel, abgesehen von Internetfähigen Geräten, sind zugelassen
 - Hilfsmittel müssen auf Ihren Tisch passen und die Stapel dürfen nicht höher als 10 cm sein
 - Handschriftliche Notizen, Bücher, Skripte, Aufgabenblätter, Musterlösungen, ...
- Es gibt keine Probeklausur
- Alle Themen sind Prüfungsrelevant (abgesehen von denen, die ich in der Vorlesung nicht behandelt habe)
- **WICHTIG: Anmeldung über CAMS-EXA/PSSO**

Das war's

- <https://aivillage.de/veranstaltungen-ai-village/>

„Film- und Diskussionsabend: KI, Bewusstsein und die Zukunft des Menschseins“

Was passiert, wenn Maschinen nicht mehr nur Werkzeuge sind? Gemeinsam schauen wir den Science-Fiction-Meilenstein Blade Runner: The Final Cut und diskutieren anschließend mit Toni Loh über die großen ethischen Fragen der KI-Zukunft: Können künstliche Wesen moralische Rechte besitzen und welche Verantwortung tragen wir gegenüber unseren... ..“

- Viel Erfolg bei den Klausuren!