



**Code of Conduct
Demokratische KI**

White Paper

Freiheit als zentraler Wert für den Einsatz von KI in der Zivilgesellschaft



White Paper

Freiheit als zentraler Wert für den Einsatz von KI in der Zivilgesellschaft

Inhalt

Einleitung.....	2
KI und Freiheit im zivilgesellschaftlichen Kontext	3
KI-Wertekompass Zivilgesellschaft	7
Was bietet der KI-Wertekompass Zivilgesellschaft?	7
Anwendungsbeispiele	12
1. Chatbot für ältere Menschen	12
2. Vergleichbarkeit von Angeboten	14
Wie geht's weiter?	16
Impressum	17



01 Einleitung

Künstliche Intelligenz (KI) wird zunehmend auch durch die Zivilgesellschaft – Organisationen, die gemeinnützig oder gemeinwohlorientiert arbeiten – eingesetzt. Eine Befragung der Universität Regensburg zeigt, dass fast 60% von 316 befragten deutschen Nichtregierungsorganisationen zumindest gelegentlich KI nutzen.¹ Doch während KI-Anwendungen zunehmend Aufgaben übernehmen, die bisher menschliche Intelligenz erforderten, birgt diese Entwicklung erhebliche Risiken für die Gesellschaft: KI-Systeme verbrauchen viel Energie und ihre Entscheidungsprozesse sind oft intransparent (sogenannte Blackbox-Effekte). Es besteht das Risiko von Datenmissbrauch sowie die Gefahr, dass sie falsche Ergebnisse erzeugen oder Fehlinformationen generieren. Solche Risiken bedrohen sowohl persönliche Autonomie als auch kollektive Freiheiten. Aufmerksamkeitszyklen, von manchen auch als „KI-Hype“ bezeichnet, können dazu führen, dass KI-Anwendungen zur vermeintlich einfachen Lösung gesellschaftlicher Probleme eingesetzt werden, ohne dass potenzielle negative Effekte ausreichend berücksichtigt werden. Umso wichtiger ist es, dass die Zivilgesellschaft selbst einen reflektierten und verantwortungsvollen Umgang mit KI-Systemen entwickelt. Insbesondere viele kleinere Organisationen, die KI einsetzen, haben bisher keine Richt- oder Leitlinien für deren Einsatz, wie eine Studie der Joseph Rowntree Foundation aus Großbritannien zeigt.² Daher ist es entscheidend, dass die Zivilgesellschaft diese (technologische) Transformation versteht und kritisch begleitet und so

Was bietet das White Paper KI und Freiheit?

- **Bewusstsein schaffen:** Das White Paper regt an, Potenziale und Risiken des Einsatzes von KI für individuelle und kollektive Freiheit zu reflektieren. Es stellt verschiedene Freiheitsdimensionen vor und zeigt Spannungen und Handlungsräume auf.
- **KI-Wertekompass:** Ein zentrales Element des White Papers ist der KI-Wertekompass für die Zivilgesellschaft. Der Kompass bietet Organisationen ein methodisches Tool, um über eigene Schwerpunkte und Möglichkeiten zu reflektieren. Der Kompass hilft außerdem zu identifizieren, welche Werte priorisiert und in Zukunft gestärkt werden sollten
- **Praxisbeispiele:** Das White Paper stellt konkrete Beispiele vor, die zeigen, welche Spannungen und Handlungsräume beim gemeinwohlorientierten Einsatz von KI entstehen.

¹ Hauer, F., Haug, S., Scharf, A., Schultz, M., & Weber, K. (2024). 3. Arbeitspapier: Künstliche Intelligenz für Nichtregierungsorganisationen – Quantitative Erforschung der Umsetzung in deutschen NROs. OTH Regensburg. DOI:[10.13140/RG.2.2.11094.66883](https://doi.org/10.13140/RG.2.2.11094.66883)

² Guler, G., Remfry, E., Kherroubi Garcia, I., Barrow, N., Duarte, T. and Ibison, Y. (2024) Grassroots and non-profit perspectives on generative AI. York: Joseph Rowntree Foundation. <https://www.jrf.org.uk/ai-for-public-good/grassroots-and-non-profit-perspectives-on-generative-ai>



ermächtigt wird, die Potenziale von KI verantwortungsbewusst für ihre Zwecke zu nutzen. Dieses White Paper beleuchtet den Grundwert Freiheit als einen zentralen Ankerpunkt für den gesellschaftlich verantwortungsvollen Einsatz von KI.

In ihrer Funktion als gesellschaftliches Vorbild und durch ihre Ausrichtung aufs Gemeinwohl hat die Zivilgesellschaft eine besondere Verantwortung, ein ausgewogenes Verhältnis zwischen Risiken und Nutzen zu finden: Während KI das Potenzial hat, menschliche Freiheit zu erweitern und zu unterstützen, besteht zugleich die Gefahr, dass sie diese untergräbt oder dominiert. Eine reflektierte und gemeinwohlorientierte Anwendung von KI kann die Zivilgesellschaft stärken, indem sie menschliche Handlungsräume erweitert. Dafür sind nicht nur klare Regeln und Kontrollmechanismen unerlässlich, sondern auch eine kontinuierliche Auseinandersetzung mit den Risiken und Folgen der Anwendung von KI. Indem es das Bewusstsein für Handlungsräume schärft und praxisnahe Beispiele liefert, hilft das White Paper, die informierte und verantwortungsvolle Nutzung von KI zu fördern.

KI und Freiheit im zivilgesellschaftlichen Kontext

Was bedeutet KI?

Unser Verständnis von KI orientiert sich an der UNESCO. Diese hat KI-Systeme wie folgt definiert:

Systeme der Künstlichen Intelligenz sind in der Lage, Daten und Informationen auf eine Weise zu verarbeiten, die menschlichen Denkprozessen ähnelt oder diese nachzuahmen scheint. Dies beinhaltet Aspekte des Denkens und Lernens, der Wahrnehmung, Vorhersage, Planung oder Steuerung.³

Wie kann KI Freiheiten schaffen?

Für den Einsatz in zivilgesellschaftlichem Kontext ergeben sich durch lernende Algorithmen, die Daten analysieren, interpretieren und darauf basierend Empfehlungen geben können, zahlreiche Anwendungsmöglichkeiten.

Beispiele für Anwendungsmöglichkeiten

- **Verständnis erleichtern:** KI-basierte Lernplattformen können spielerisch und personalisiert komplexe Herausforderungen interaktiv erklären und so der Zivilgesellschaft helfen, mit einem breiteren Publikum zu kommunizieren.
- **Prozesse vereinfachen:** KI-Anwendungen können bei komplexen Antragsprozessen helfen, Fragen von Spender:innen beantworten oder Ehrenamtliche vernetzen und so Zeit für kreative und menschenzentrierte Arbeit schaffen.
- **Prognosen entwickeln:** KI-Systeme können durch die Verarbeitung großer Datenmengen Trends analysieren oder Szenarien simulieren – ob im Spendenverhalten oder in der Klimafolgenabschätzung. So können Organisationen proaktiv auf Entwicklungen reagieren und fundierte Entscheidungen treffen.

³ UNESCO (2023): Ethik der künstlichen Intelligenz: Wegweiser für die Gestaltung unserer Zukunft. S.8. https://www.unesco.de/sites/default/files/2023-05/DUK_Broschuere_KI_A5_DE_web.pdf



Freiheit im Kontext von KI

Freiheit beschreibt die Fähigkeit und das Recht eines Individuums, ohne äußeren Zwang zu handeln, eigene Entscheidungen zu treffen und Lebensumstände selbst zu bestimmen. Freiheit ermöglicht es, individuelle Potenziale zu entfalten und Neues auszuprobieren. Entscheidend dafür sind Urteilsfähigkeit und Möglichkeiten, zwischen echten Alternativen wählen zu können. Dazu gehört auch die Möglichkeit, sich bewusst gegen die Nutzung von KI zu

entscheiden oder alternative, ressourcenschonende Lösungen zu wählen. Damit ist Freiheit ein zentraler Wert, um sicherzustellen, dass KI einen echten Nutzen für Menschen erbringt.

Was ist Freiheit?

- Selbstbestimmung, Selbstbestimmtheit, Potenzialnutzung
- Wahlfreiheit im Handeln und Entscheiden
- Abwesenheit von Zwang

Was braucht es, um Freiheit zu schaffen?

Transparenz und Kontrolle gewährleisten

- KI-Systeme sollten so transparent gestaltet sein, dass ihre Funktionsweise für die Nutzer:innen nachvollziehbar ist. Zusätzlich sollten umfassende Kontrollmöglichkeiten vorhanden sein, um ihre Auswirkungen steuern zu können.

Selbstbestimmte KI-Nutzung ermöglichen

- Um echte Autonomie und selbstbestimmte Entscheidungen zu ermöglichen, ist der Zugang zu qualitativ hochwertiger Bildung unerlässlich. Mitarbeitende und Betroffene brauchen Schulungen und Weiterbildungen, um eine selbstbestimmte, verantwortungsvolle und aufgeklärte Nutzung von KI zu gewährleisten.

Grenzen von Freiheit anerkennen

- Freiheit endet dort, wo sie die Freiheit anderer einschränkt. Sie umfasst sowohl individuelle als auch gesellschaftliche Verantwortung. Freiheit muss daher innerhalb gesellschaftlicher und rechtlicher Grenzen ausgeübt werden, wie etwa dem Schutz der Privatsphäre, der Vermeidung von Diskriminierung und dem bewussten Umgang mit Ressourcen.



Was müssen wir beim Einsatz von KI beachten?

Risiken beim Einsatz von KI⁴

- **Autonomie und Entscheidungsfreiheit:** KI-Systeme können individuelle Entscheidungsfreiheit einschränken, indem sie Entscheidungen durch vorgegebene Algorithmen und auf Basis verzerrter Datengrundlagen beeinflussen. Die zunehmende Automatisierung durch selbstlernende Systeme birgt das Risiko, die Kontrolle über Entscheidungsprozesse zu verlieren.
- **Meinungsfreiheit und Fehlinformationen:** KI-Systeme können die Verbreitung von Desinformationen erleichtern, indem Inhalte schnell erstellt, manipuliert und angepasst werden können. Da KI-Systeme häufig als objektiv und neutral wahrgenommen werden, besteht zudem die Gefahr, dass Fehler, Verzerrungen oder Halluzinationen übersehen werden, und so falsche Informationen aus Versehen verbreitet werden. KI spielt außerdem eine wichtige Rolle in der Moderation von Inhalten und kann hier einerseits menschliche Fehleranfälligkeit verringern, andererseits aber auch bei mangelnder menschlicher Aufsicht die Meinungsfreiheit einschränken.
- **Rechtliche Risiken und Urheberrecht:** KI kann die Erstellung und Verbreitung schädlicher oder strafrechtlich relevanter Inhalte, wie Deepfakes oder Schadsoftware, erleichtern. Dieses Missbrauchspotenzial birgt erhebliche gesellschaftliche und rechtliche Risiken. Gleichzeitig besteht bei der Nutzung von Sprachmodellen weiterhin Unsicherheit im Bereich geistiges Eigentum und Urheberrecht.
- **Diskriminierung und Ungleichheit:** KI-Systeme können bestehende Diskriminierungen und Ungleichheiten verschärfen, indem sie beispielsweise diskriminierende Muster aus Trainingsdaten übernehmen.
- **Datenschutz und Überwachung:** Durch fehlende Einwilligung zur Datenverarbeitung, Datenmissbrauch oder KI-gestützte Überwachung können Datenschutz, Persönlichkeitsrechte und Privatsphäre bedroht sein.
- **Ressourcenverbrauch und Rebound-Effekte:** Der Einsatz von KI kann durch Effizienzsteigerungen den Ressourcenverbrauch senken, jedoch auch zu sogenannten Rebound-Effekten führen, bei denen Gesamtverbrauch und Umweltbelastungen durch erhöhte Nutzung letztlich wieder ansteigen, beispielsweise durch den Stromverbrauch von Rechenzentren.
- **Technologische Abhängigkeiten:** Die Nutzung von KI-Anwendungen kann zu Abhängigkeiten von großen KI-Anbietern führen und die Wahlfreiheit einschränken, insbesondere wenn Organisationen der Zugang zu alternativen Anwendungen fehlt. Zudem kann die Automatisierung wichtiger Aufgaben langfristig zum Verlust wichtiger menschlicher Kompetenzen führen.

⁴ Siehe Council of Europe (2023) Study on the impact of artificial intelligence systems, their potential for promoting equality, including gender equality, and the risks they may cause in relation to non-discrimination. <https://rm.coe.int/prems-112923-qbr-2530-etude-sur-l-impact-de-ai-web-a5-1-2788-3289-7544/1680ac7936>; Deutscher Ethikrat (2023) Stellungnahme: Mensch und Maschine – Herausforderungen durch Künstliche Intelligenz. <https://www.ethikrat.org/fileadmin/Publikationen/Stellungnahmen/deutsch/stellungnahme-mensch-und-maschine.pdf>; Deutscher Bundestag (2020) Bericht der Enquete-Kommission Künstliche Intelligenz – Gesellschaftliche Verantwortung und wirtschaftliche, soziale und ökologische Potenziale Drucksache. 19/23700. <https://dserver.bundestag.de/btd/19/237/1923700.pdf>



Position und Handlungsräume der Zivilgesellschaft verstehen

Um nachhaltige Abwägungen treffen zu können und sich selbst als gestaltend zu begreifen, müssen zivilgesellschaftliche Organisationen eine klare Idee ihrer Handlungsräume entwickeln. Dies umfasst mehrere Aspekte:

1. Bewusstsein für die Verfestigung problematischer Strukturen: Organisationen sollten bedenken, dass die Nutzung von Anwendungen großer Anbieter Monopole stärken oder Abhängigkeiten schaffen kann.
2. Anerkennung struktureller Zwänge: Besonders kleinere oder technisch weniger versierte Organisationen haben oft begrenzte Ressourcen. Daher ist es wichtig, einen offenen und ehrlichen Austausch über wahrgenommene Möglichkeiten zu fördern.
3. Respekt für Diversität in Meinungen und Verhaltensweisen: Zivilgesellschaftliche Organisationen haben unterschiedliche Werte und Schwerpunkte, die es zu respektieren gilt. Damit Organisationen ihre eigene Positionierung reflektieren können, stellen wir den KI-Wertekompass vor.



02 KI-Wertekompass Zivilgesellschaft

Der *KI-Wertekompass Zivilgesellschaft* bietet zivilgesellschaftlichen Akteuren ein umfassendes Instrument zur Reflexion ihrer Positionen, Strategien und Herausforderungen im Bereich der gemeinwohlorientierten Nutzung von KI. Er beschreibt und visualisiert die Zusammenhänge und Spannungsfelder von KI und Freiheit und dient als Denkwerkzeug, um verschiedene Dimensionen von Freiheit zu erfassen und Handlungsräume als Freiräume zu erkennen. Als Visualisierungstool unterstützt der KI-Wertekompass den Dialog mit Beteiligten und Betroffenen.

Was bietet der KI-Wertekompass Zivilgesellschaft?

Der Kompass zeigt verschiedene Freiheitsdimensionen auf, die je nach Kontext unterschiedlich priorisiert und interpretiert werden können. Zivilgesellschaftliche Organisationen bewegen sich zum Beispiel oft zwischen den Polen Offenheit (z.B. Transparenz, Open Data) und Sicherheit (z.B. Datenschutz, Privatsphäre): So sind Transparenz und Open Data beim Einsatz von KI zur Analyse von Klimadaten zentral, während bei der Verarbeitung sensibler Daten von vulnerablen Gruppen die Priorität auf Sicherheit liegen sollte. Organisationen und Individuen legen ihre Position basierend auf ihren Werten, Zielen und dem spezifischen Kontext fest. Die Verwendung des Wertekompasses fördert ein tieferes Verständnis der komplexen Zusammenhänge und ermöglicht fundierte Entscheidungen im Bereich KI und digitaler Transformation.

Dimensionen von Freiheit

Freiheit kann philosophisch, politisch, sozial, individuell, gesellschaftlich, positiv oder negativ verstanden werden. Statt vermeintlich gegensätzliche Freiheitsverständnisse jedoch als unüberwindbare Hindernisse zu sehen, können wir sie als Spektren von Möglichkeiten betrachten. Jede Ausprägung repräsentiert ein spezifisches Verständnis von Freiheit, die auch miteinander in Verbindung stehen. Jede Position auf dem Kompass steht für einen potenziellen Handlungsraum mit eigenen Vor- und Nachteilen.

Offenheit vs. Sicherheit

Offenheit: Diese Ausprägung priorisiert die Transparenz, Nachvollziehbarkeit und Partizipationsmöglichkeiten von Prozessen. Dies umfasst die Erklärung und Erklärbarkeit von Modellen, den verwendeten Algorithmen und Methoden oder die Veröffentlichung der Datengrundlage von KI-Systemen.

Beispiel: Eine NGO will mithilfe von KI die CO₂-Emissionen des Verkehrs in einer Stadt messen. Der Prozess wird möglichst partizipativ gestaltet und sowohl die Trainingsdaten als auch die Ergebnisse werden auf einer Website der Öffentlichkeit zur Verfügung gestellt.



Beispiel: Eine Genossenschaft nutzt ein selbstlernendes KI-System, um Netzwerkaktivitäten zu überwachen und auf mögliche Cyberattacken zu reagieren. Die Genossenschaft verzichtet auf Offenlegung, um Missbrauch zu verhindern.

Sicherheit: Bei dieser Ausprägung geht es um den Schutz vor unbefugtem Zugriff und Nutzung, beispielsweise um sicherzustellen, dass sensible Daten geschützt sind oder KI-Systeme nicht manipuliert oder gehackt werden können. Es kann dabei auch um den Schutz von Geschäftsgeheimnissen oder proprietärer Technologien gehen.

Explorativ vs. Konservativ

Explorativ: Diese Ausprägung beschreibt das Erproben neuer Technologien und Ansätze, um innovative Lösungen und Ideen zu erzielen. Dies könnte die Anwendung von KI in neuen Bereichen oder die Entwicklung neuer, disruptiver Anwendungen umfassen.

Beispiel: Eine Selbsthilfeorganisation testet in einem Pilotprojekt KI-Sprachassistenten, um Menschen mit Behinderungen in ihrer selbstständigen Lebensführung zu unterstützen.

Beispiel: In einer Wohlfahrtsorganisation bleibt der Einstellungsprozess durch Führungskräfte bestehen. Nach Abschluss der Personalentscheidungen analysiert jedoch ein KI-System die Entscheidungen auf mögliche diskriminierende Muster, um den Führungskräften vertrauliches Feedback zu ihrer Einstellungspraxis zu geben.

Konservativ: Diese Ausprägung fokussiert sich auf das Beibehalten bewährter Praktiken, um Stabilität und Vorhersehbarkeit zu gewährleisten. Dies kann bedeuten, dass neue Technologien sorgfältig und schrittweise eingeführt werden oder auch auf experimentelle oder kaum nachvollziehbare KI-Anwendungen verzichtet wird.

KI-Minimalismus vs. KI-Maximierung

KI-Minimalismus: Hierbei werden der Einsatz von KI bewusst auf das Notwendige beschränkt oder KI-Methoden eingesetzt, die wenig Rechenleistung erfordern und eine durchschaubare Funktionsweise haben. So werden Freiheit von technologischen Abhängigkeiten, geringerer Ressourcenverbrauch und hohe menschliche Kontrolle gefördert.

Beispiel: Eine Bürgerinitiative nutzt Checklisten und Leitfäden anstelle von KI-gestützten Schreibassistenten, um die Qualität der Social-Media- und Blogbeiträge ihrer Ehrenamtlichen zu verbessern.

Beispiel: Eine Umweltstiftung nutzt Drohnen mit KI-Bilderkennung sowie automatische Datenanalysen und -visualisierungen, um die Datengrundlage zur Biodiversität zu verbessern.

KI-Maximierung: Hier wird die maximale Nutzung von KI und datengetriebener Prozesse zur Erweiterung menschlicher Fähigkeiten angestrebt, sodass alle technischen Möglichkeiten ausgeschöpft werden.



Menschliche Autor:innenschaft vs. KI-generierte Inhalte

Menschliche Autor:innenschaft: Hier geht es um menschliche Kreativität und Kontrolle in der Erstellung von Inhalten. Dabei steht die Förderung individueller Entscheidungsfreiheit und Autonomie im Fokus.

Beispiel: Zum Tracking globaler Regulierungstrends zum Datenschutz stellt eine gemeinnützige GmbH auf Basis einer Datenbank zu nationalen Gesetzesvorhaben automatisiert kurze Länderreports und Visualisierungen auf ihrer Website zur Verfügung.

Beispiel: Die investigative Analyse einer NGO zu lokaler Korruption wird durch ein KI-System auf sprachliche Verständlichkeit und grammatikalische Fehler geprüft. Um die Ergebnisse möglichst präzise wiederzugeben und rechtssichere Formulierungen beizubehalten, wird jedoch die menschliche Autor:innenschaft priorisiert.

KI-generierte Inhalte: Hierbei wird KI zur Erstellung von Inhalten genutzt und gestaltet die Inhalte weitgehend selbstständig. Ziel ist es, KI für gemeinschaftliche Problemanalysen und -lösungen einzusetzen.

Präzision vs. Imagination

Präzision: Bei dieser Ausprägung liegt der Fokus auf Präzision und Zuverlässigkeit von KI-Systemen, um möglichst genaue und konsistente Ergebnisse zu gewährleisten.

Beispiel: In einem Workshop nutzt eine Einrichtung, die sich für die Gleichstellung der Geschlechter einsetzt, einen KI-Bildgenerator, der utopische Zukünfte visualisiert, um möglichst kreative Antworten zu erhalten.

Beispiel: Beim Design eines Chatbots, der Informationen für Betroffene von Diskriminierung oder Gewalt liefert, schränkt ein Social Startup die möglichen Aussagen stark ein, um fehlerfreie und präzise Antworten zu gewährleisten.

Imagination: Hierbei liegt ein Schwerpunkt auf der Förderung innovativer und unerwarteter Lösungen durch KI, auch wenn diese zum Teil weniger präzise oder faktentreu sind.

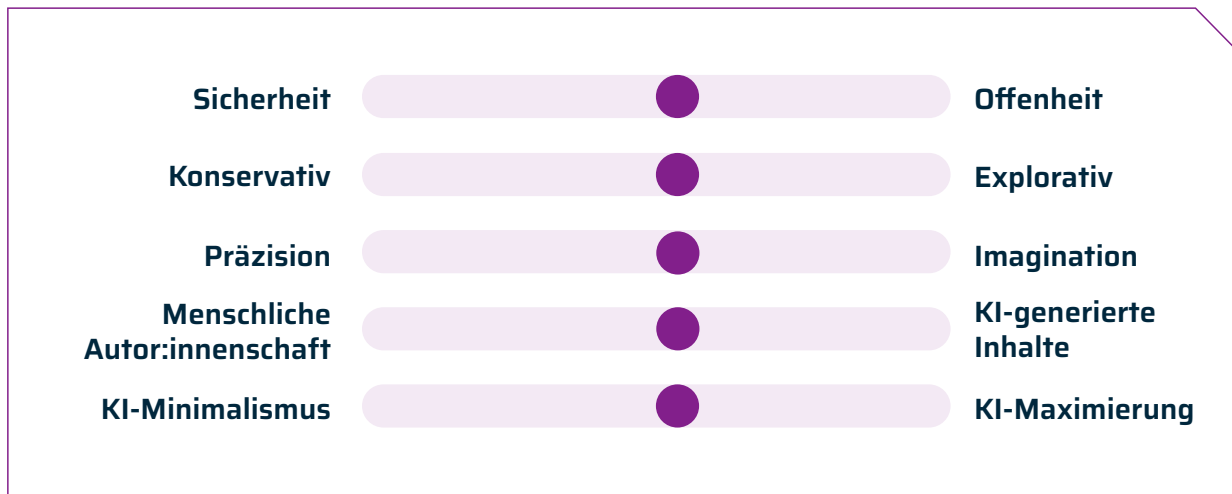


Abb. 1 Visualisierung der Dimensionen des KI-Wertekompasses Zivilgesellschaft

Wie setze ich den Kompass ein?

- 1. Positionsbestimmung:** Organisationen können ihre aktuelle sowie angestrebte Position auf dem Kompass markieren, um über ihre Rolle zu reflektieren. Sie können beispielsweise entscheiden, ob sie in einem bestimmten Projektkontext eher explorativ (neue KI-Technologien ausprobieren) oder konservativ (bestehende Praktiken bewahren) agieren (möchten).
- 2. Risiko- und Chancenbewertung:** Der Kompass bietet durch die Verortung auch einen Analyserahmen, um Chancen und Risiken zu bewerten. Durch die verschiedenen Dimensionen können Projekte ganzheitlich beurteilt werden.
- 3. Strategieentwicklung:** Mithilfe der Positionierung und -bewertung können Organisationen die Auswirkungen unterschiedlicher Entscheidungen visualisieren und Strategien entwickeln, um sowohl die Chancen der KI zu nutzen als auch die Werte ihrer Organisation zu bewahren. Das ermöglicht sowohl interne Diskussionen über ethische Fragestellungen als auch die Profilschärfung nach außen.



Was ist „richtig“?

Die „richtige“ Position auf jeder Achse hängt stark vom spezifischen Kontext, den verfügbaren Ressourcen, der Mission und den Werten der jeweiligen Organisation ab. Je nach Projekt, Kontext und Herausforderungen können Organisationen ihre Position anpassen und zwischen verschiedenen Ausprägungen wechseln. Dabei müssen verschiedene Faktoren sorgfältig gegeneinander abgewogen werden. Dazu gehören Nachhaltigkeit und Ressourcenschonung, Risiko- und Technikfolgenabschätzung, Wissen und Kompetenzen. Jede Positionierung hat Konsequenzen, die sorgfältig reflektiert werden müssen. Und dies nicht nur im Kontext von Freiheit, sondern auch im Zusammenhang mit anderen Grundwerten wie Gerechtigkeit und Solidarität.

KI-Wertekompass als Entscheidungshilfe

Der KI-Wertekompass kann zivilgesellschaftlichen Organisationen dabei helfen, unter Berücksichtigung ihrer Werte und Ressourcen Entscheidungen zu fällen. Die Abwägung zwischen minimalistischer und maximalistischer oder konservativer und explorativer Anwendung von KI kann zum Beispiel die Verteilung von Ressourcen beeinflussen. Wo lohnen sich Investitionen in KI-Kapazitäten und wann ist ein alternativer, ressourcenschonender Ansatz sinnvoller? Die Dimension Offenheit vs. Sicherheit kann dagegen die Grenzen eigener Handlungsfähigkeit aufzeigen: Welchen Risiken kann man selbst vorbeugen, wo braucht es Änderungen auf Ebene der Betreiber:innen oder in der Politik?



03 Anwendungsbeispiele

1. Chatbot für ältere Menschen

Kontext und Herausforderung

Ein Sozialverband bietet älteren Menschen regelmäßige Gesprächsangebote, wie etwa Telefon- und Besuchsdienste. Trotzdem gibt es viele Senior:innen, die isoliert oder einsam sind und einen Bedarf an Austausch haben, der durch Mitarbeitende wegen Zeit- und Ressourcenmangel nicht gedeckt werden kann. Daher entsteht die Idee, eine Anwendung, die natürlichen sprachlichen Dialog mit einem technischen System ermöglicht, einzurichten, um den menschlichen Kontakt zu ergänzen. Der KI-basierte Chatbot sollte einen freundlichen und verständlichen Austausch ermöglichen, gleichzeitig dürfen die Nutzer:innen keine unrealistischen Erwartungen entwickeln. Nach Abwägung der möglichen Risiken und Chancen wird der KI-Wertekompass genutzt, um die Entscheidung für die Gestaltung eines KI-Tools entlang wichtiger Werte und Kriterien zu steuern.

Offenheit vs. Sicherheit

Der Einsatz des Chatbots sollte transparent sein, insbesondere in Bezug auf die Fähigkeiten und Grenzen des Tools. Es sollte klar kommuniziert werden, dass es sich um einen digitalen Assistenten handelt, der echte menschliche Interaktion, zum Beispiel mit Ärzt:innen oder Familie, nicht ersetzen kann. Die Entwicklung des Chatbots sollte gemeinsam mit den älteren Menschen erfolgen, um eine möglichst bedarfsgerechte Gestaltung zu gewährleisten. Gleichzeitig ist der Schutz der persönlichen Daten der Nutzer:innen zentral. Auch Trainingsdaten können sensible Informationen zur Gesundheit der Nutzer:innen enthalten und sollten nicht öffentlich zugänglich gemacht werden.

Explorativ vs. Konservativ

Der Chatbot sollte menschliche Interaktion ergänzen. In einer Pilotphase könnten mit einer Kleingruppe verschiedene Gesprächsthemen und -stile begleitet getestet werden, um den Chatbot möglichst zugänglich zu gestalten. Nach ausführlichen Tests könnten die Anwendungsfälle ausgeweitet werden, um Ältere beispielsweise auch bei Alltagsproblemen oder Verwaltungsaufgaben zu unterstützen. Ein konservativerer Ansatz würde den Chatbot nur als zusätzliche Unterstützung während der Anwesenheit anderer Personen zum Einsatz kommen lassen.

KI-Minimalismus vs. KI-Maximierung

Ein minimalistischer Chatbot könnte beispielsweise eine Suchmaschine ergänzen, um Rezepte, Fernsehsendungen oder bestimmte Informationen leichter auffindbar zu machen. Eine weiter entwickelte Option würde Gesprächsangebote zu komplexen Themen, wie Politik, Kunst und Kultur oder sensiblen Themen wie Trauer, Depression machen und sich auf die Stimmung, Vorlieben und Interessen der Nutzer:innen einstellen.



Menschliche Autor:innenschaft vs. KI-generierte Inhalte

Um die menschliche Autor:innenschaft zu stärken, könnten die Inhalte durch die Älteren selbst, Familie und Pflege- und Betreuungspersonal eng vorgegeben werden und sich auf bestehende Veranstaltungen, Protokolle oder aufgezeichnete Unterhaltungen stützen. Im Gegensatz dazu könnten KI-generierte Inhalte unabhängig von diesen Vorgaben dynamisch auf die Wünsche der Nutzer:innen reagieren. Dadurch würde der Chatbot eigene Inhalte wie Bilder, Geschichten oder neues Wissen schaffen.

Präzision vs. Imagination

Während bei Auskünften zu Krankheiten oder Verwaltungsaufgaben präzise Antworten priorisiert würden, könnte ein kreativer Chatbot Spiele erfinden, Empfehlungen für Hobbys geben oder Konversationen zu neuen Themen anstoßen.

Fazit

Der Sozialverband entscheidet sich, in einer kleinen Gruppe von geschulter Nutzer:innen den KI-basierten Chatbot explorativ zu testen, der sich stark an den Wünschen der Nutzer:innen orientiert, und daher spielerisch sowie kreativ gestaltet ist, um neue Impulse zu setzen. Gespräche werden aufgezeichnet und für das Training des Chatbots lokal gespeichert, können jedoch nicht für weitere Zwecke genutzt werden.

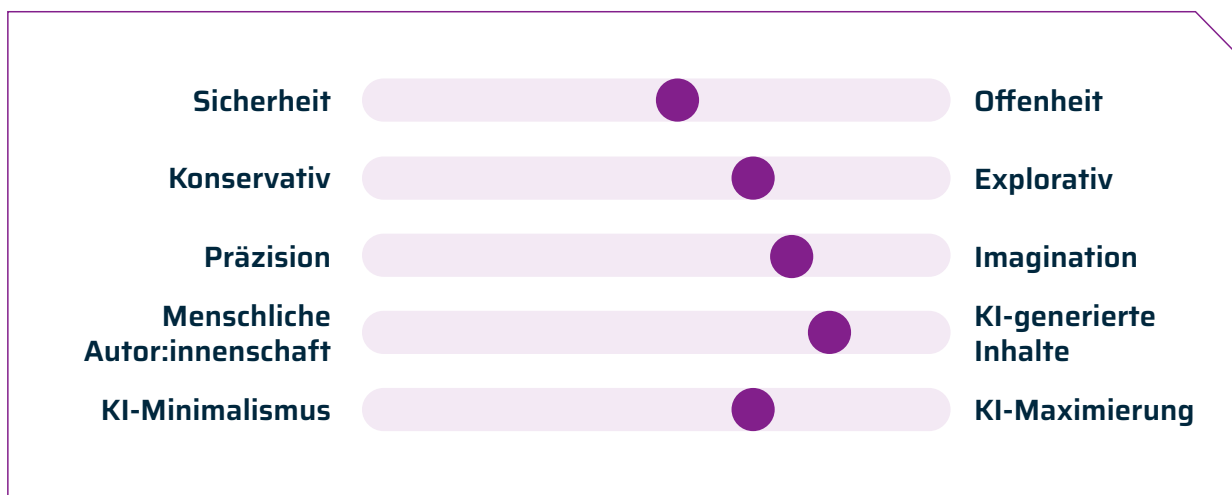


Abb. 2 Visualisierung der Ergebnisse zum Chatbot



2. Vergleichbarkeit von Angeboten

Kontext und Herausforderung

Eine NGO holt verschiedene Angebote für die Erstellung und langfristige Betreuung einer neuen Website ein. Die Angebote betonen jedoch unterschiedliche Aspekte, wie Support- und Backupmöglichkeiten, Mindestvertragslaufzeiten, Datenschutzinformationen oder Marketingtools, was die direkte Vergleichbarkeit erschwert.

Aktuell prüft das Team die Angebote in einer Tabelle. Durch die Vielfalt der Kriterien sowie die technische Komplexität ist der Prozess mühsam und zeitintensiv. Außerdem fehlen standardisierte Bewertungsmaßstäbe, um die Angebote objektiv zu vergleichen. Nach Abwägung der möglichen Risiken und Chancen wird der KI-Wertekompass genutzt, um die Entscheidung über den Einsatz eines KI-basierten Textanalysetools entlang wichtiger Werte und Kriterien zu steuern.

Offenheit. vs. Sicherheit

Die Nachvollziehbarkeit und Transparenz der Angebotswahl kann durch den Einsatz einer KI-Anwendung, die Informationen aus Dokumenten extrahieren und nach verschiedenen Kriterien bewerten kann, idealerweise erhöht werden, zum Beispiel durch klare Vergleichskriterien. Angebote sind jedoch häufig vertraulich und können persönliche Daten enthalten. Daher sollten Datenschutz- und Vertraulichkeitsaspekte bei der Wahl des Tools und während der Durchführung berücksichtigt werden.

Explorativ vs. Konservativ

Die KI-Anwendung wird zunächst parallel zum menschlichen Vergleichsprozess getestet, um mögliche Schwächen zu identifizieren. Danach ist ein eher explorativer Ansatz vorgesehen, um zentrale Kriterien und deren Gewichtung zu ermitteln. Über den Vergleich hinaus könnte ein KI-Tool eingesetzt werden, um mögliche Szenarien und deren Auswirkungen (wie verstärkte Cyberangriffe oder plötzlich stark ansteigende Zugriffszahlen) bei verschiedenen Anbietern zu simulieren.

KI-Minimalismus vs. KI-Maximierung

Ein minimalistischer Ansatz könnte bedeuten, KI-Tools lediglich zur Verbesserung des manuellen Vergleichs zu nutzen, beispielsweise durch Plugins, die natürliche Sprache in Formeln umwandeln. Das Team wünscht sich jedoch eine umfassende Analyse mit Vorschlägen zur Gewichtung verschiedener Kriterien. Ein maximalistischer Ansatz könnte die automatisierte Entscheidung einer KI-Anwendung beinhalten und gegebenenfalls auch die Kommunikation mit den Anbietern an einen Chatbot auslagern.

Menschliche Autor:innenschaft vs. KI-generierte Inhalte

Die KI-Anwendung soll lediglich ein Vergleichsdokument, beispielsweise eine Tabelle erstellen, könnte aber gegebenenfalls kurze Textzusammenfassungen schreiben, die Vor- und Nachteile der jeweiligen Anbieter auch für weitere Teammitglieder zusammenfassen.



Präzision vs. Imagination

Bei diesem Anwendungsfall liegt der Fokus klar auf Präzision. Falls die KI-Anwendung Inhalte erfindet, könnten Anbieter durch den Prozess unfair benachteiligt werden. Die Anwendung sollte genaue Daten auslesen und vergleichen können, um objektive Ergebnisse zu liefern.

Fazit

Eine passende KI-Anwendung, die Informationen extrahieren und nach verschiedenen Kriterien tabellarisch auswerten kann, wird ausgewählt. Das Team informiert die Anbieter bereits bei Einholung der Angebote darüber, dass diese durch eine KI-basierte Textanalysen verglichen werden. Deshalb bittet es um pseudonymisierte Versionen der Angebote. Nach einem Testlauf wird das Tool eingesetzt, um eine umfassende und leicht verständliche Analyse der Angebote mit transparenter Darstellung der einzelnen Kriterien zu erhalten. Das Team legt die Gewichtung der Kriterien selbst fest und erhält zwei Empfehlungen, die durch eine:n Mitarbeiter:in geprüft werden. Der menschliche Entscheidungsprozess bleibt jedoch entscheidend.

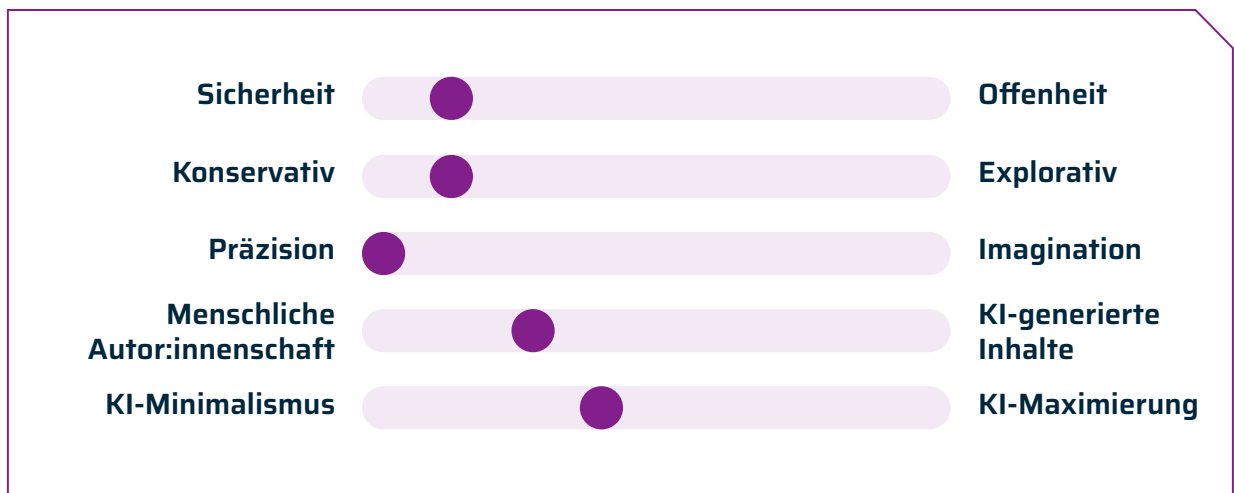


Abb. 3 Visualisierung der Ergebnisse zum Vergleich von Angeboten



Wie geht's weiter?

Dieses White Paper ist nur der Anfang. Jetzt seid ihr gefragt: Testet den KI-Wertekompass und lasst uns wissen, wie ihr die verschiedenen Dimensionen bewertet. Was ist für eure Organisation entscheidend, um KI gemeinwohlorientiert zu nutzen? Wie schafft ihr es, dass KI eure Handlungsfreiheit erweitert statt einschränkt? Eure Erfahrungen, Ideen und Anregungen sind für uns von großer Bedeutung. Schreibt uns euer Feedback und eure Positionen an: buero@d-64.org



Dieses Positionspapier ist in einem internen Arbeitsprozess des Projekts **Code of Conduct Demokratische KI** unter der Leitung von **Anke Obendiek**, PhD (D64) entstanden. Mitarbeit: **Ulrich Berger** (D64), **Ruven Börger** (ASB Deutschland e.V.), **Torben Dzillak** (D64), **Jonas Ferdinand** (D64), **Michael Golze** (Deutsche Umweltstiftung), **Lou Heidrich** (Deutsches Rotes Kreuz), **Niels Heinemann** (Kompetenzzentrum Technik - Diversity - Chancengleichheit e. V.), **Monika Ilves** (D64), **Barbara Jung** (AKTIVOLI Landesnetzwerk Hamburg e.V.), **Maximilian Kühn** (AWO Bundesverband e.V.), **Nevena Nikolajevic** (Correlaid, e.V.), **Damian Paderta** (Offene Kommunen. NRW), **Oliver Pfeiderer** (LAG Selbsthilfe RLP), **Moritz Ritter** (Liquid Democracy e.V.), **Rebecca Roth** (Neue deutsche Medienmacher*innen), **Tim Schrock** (Arbeitskreis deutscher Bildungsstätten e.V. (AdB)), **Jonas Stettner** (Correlaid, e.V.), **Sabine Wolf** (BAGSO - Bundesarbeitsgemeinschaft der Seniorenorganisationen e.V.) und **Sibylle Würz** (FrauenComputerZentrumBerlin e.V. (FCZB)).



Impressum

Wir bedanken uns bei allen, die uns mit ihren Hinweisen und Kommentaren bei der Erstellung des Papiers weitergeholfen haben.

Gestaltungskonzept: **Antje Schnier**

Gestaltung: D64 / **Johann Lensing**

Adresse

D64 – Zentrum für Digitalen Fortschritt e.V.

Co-Vorsitzender: Erik Tuchtfeld

Co-Vorsitzende: Svea Windwehr

Chausseestraße 5

10115 Berlin

Kontakt

Anke Obendiek, PhD (D64 – Zentrum für Digitalen Fortschritt)

buero@d-64.org

demokratische-ki.de

November 2024

Gefördert vom:



Bundesministerium
für Familie, Senioren, Frauen
und Jugend