

Ethik, Governance und Regulierung von KI-Systemen: Eine ungeklärte Machtfrage

Wie lassen sich menschliche Grundwerte – von Inklusion und Transparenz bis hin zu Gerechtigkeit und Menschenwürde – in der Entwicklung und Nutzung von KI-Systemen verankern? Und wie stellen wir sicher, dass KI-Systeme nicht Märkte, Informationsflüsse und Staaten kontrollieren, sondern im Dienst von Demokratie, Nachhaltigkeit und Gemeinwohl stehen? Auf diese komplexen Fragen gibt es keine einfachen Antworten. Klar ist jedoch: Eine starke Governance und ein verlässlicher regulatorischer Rahmen zählen zu den zentralen Schlüsseln.



Von Luisa Lange, Sustainability Management, Ethius Invest

So abstrakt KI-Systeme auch erscheinen mögen, sie können sich auf die physische Welt, das soziale Miteinander, unsere Gesellschaften, Staaten und Demokratien wie auch auf das Leben Einzelner im Zweifel sehr drastisch auswirken. Beginnen wir mit konkreten Beispielen, wie KI-Systeme massiv in das Leben von Individuen eingreifen können:

- Wenn ein KI-System auf Basis vorgelegener Daten darauf trainiert wird, Bewerbungen zu analysieren, und darauf aufbauend über Einstellungen

entschieden wird, bleiben dem Unternehmen möglicherweise nicht nur die besten Fachleute vorenthalten, sondern es werden zugleich auch konkrete Bewerber und Bewerberinnen diskriminiert.

- Wenn KI-Systeme eingesetzt werden, um Sozialbetrug aufzudecken, dabei aber – wie in den Niederlanden geschehen – Grundrechte nicht eingehalten werden (Rachovitsa 2022), kann sich dies für ohnehin besonders vulnerable Menschen als sehr folgenreich erweisen.

- Wenn, wie offenbar in der Praxis bereits geschehen, in Kriegen KI-Systeme über Bombardierungsziele entscheiden, ohne dass dies von Menschen kritisch geprüft wird (Singh 2024), kann sich dies für Zivilisten und Zivilistinnen verheerend auswirken.

Indikatoren für Verantwortungsübernahme

Alle diese Beispiele zeigen, welche Bedeutung Aspekten wie Verantwortlichkeiten, Zuständigkeiten und Haftungsregelungen sowohl in der Entwicklung als auch während der Nutzung von KI-Systemen zukommt. Das IÖW (2021) hat daher neben sozialen und ökologischen Indikatoren auch für den Bereich der Governance Anforderungen entwickelt, die beispielsweise vorsehen, dass Ansprechpersonen für ethische Belange eingesetzt werden, die dann als vermittelnde Instanzen agieren, um Konflikte in Bereichen wie Transparenz, Fairness, Verantwortung, Datenschutz oder Autonomie zu regeln und zu lösen. Dies kann über einen KI-Ethikausschuss operationalisiert werden.

Rollenfragen, Dokumentation und Haftung

Ein weiterer Indikator für gute Governance ist laut IÖW (ebenda), dass Aufgaben, Rollen und Verantwortungsreiche definiert, dokumentiert und klar voneinander abgegrenzt werden – etwa zwischen einsetzender und entwickelnder Organisation. Unterstützend hierfür wirken regelmässige Updates der Dokumentation, die an sich vollständig, detailliert und transparent sein sollte.

Bei Haftungsfragen ist es darüber hinaus wichtig, für Fälle von möglichen Schäden Absicherungen durch finanzielle Rücklagen, Versicherungen und andere Kompensationsformen vorzusehen. Nicht zuletzt wird ein Code of Conduct als sinnvoll erachtet. Dieser soll Prinzipien und Standards umfassen sowie grundlegende Werte definieren, welche Organisationen, die KI-Systeme entwickeln oder nutzen, anleiten und auf ethische Zielsetzungen rückbinden.

Die Beteiligung von Anspruchsgruppen sicherstellen

Ein weiterer zentraler Faktor ist die Einbindung der potenziell von den KI-Systemen Betroffenen. Auch dieser Aspekt ist in die Governance von KI-Systemen zu integrieren (ebenda, S. 34f.). Um dies zu gewährleisten, müssen zunächst die betreffenden Anspruchsgruppen identifiziert und in einem zweiten Schritt sinnvoll klassifiziert werden. Erst dann können sie systematisch eingebunden werden, was allerdings nicht nur im Entwicklungsprozess erforderlich ist, sondern über den gesamten Lebenszyklus der KI-Systeme hinweg erfolgen sollte. Als inhaltlich wichtig für die Einbindung der Anspruchsgruppen wird genannt, dass Abläufe erklärbar sein sollten, die Autonomie des Individuums respektiert wird und Fairness ein Leitmotiv ist.

Risikomanagement

Da der Einsatz von KI-Systemen mit vielfältigen Risiken verbunden ist, sollten Unternehmen darüber hinaus ein umfassendes Risikomanagement etablieren (ebenda). Dieses umfasst insbesondere die Identifikation, Klassifizierung und Bewertung potenzieller Risiken, etwa in den Bereichen Diskriminierung, Datenschutz, IT-Sicherheit und möglicher Angriffe auf KI-Systeme. Darüber hinaus müssen geeignete Massnahmen zur Risikominimierung umgesetzt werden, beispielsweise technische Schutzmechanismen sowie ein kontinuierliches Monitoring. Zudem sollte das Risikomanagement fest in bestehende Unternehmensprozesse integriert werden – nicht zuletzt auch, um regulatorische Anforderungen und Nachhaltigkeitsziele zuverlässig sicherzustellen.

Wem KI-Systeme nutzen, ist eine Machtfrage

Damit sind wesentliche Aspekte genannt, die dazu beitragen können, dass KI-Systeme sowohl in der Entwicklung als auch in der Anwendung ethische Grundwerte berücksichtigen und Risiken für Mensch und Gesellschaft minimieren. Doch das allein greift zu kurz. KI-Systeme rühren weit darüber hinaus an ökonomische Machtfragen, können Demokratien gefährden und gewinnen in geopolitischen, strategischen Entscheidungen zunehmend an Bedeutung. Dabei liegt das Problem nicht in der Technologie selbst, sondern im Umgang mit ihr. Der deutsche Ökonom Marcel Fratzscher (2026) bringt dies wie folgt auf den Punkt: «Nicht die KI selbst bedroht die Demokratie, sondern die politische Ökonomie ihrer Entwicklung (...). Wenn sich KI-Ressourcen bei wenigen privaten Akteuren konzentrieren, dann verschiebt sich Macht aus demokratisch kontrollierten Räumen in intransparente Sphären. Wer KI kontrolliert, kontrolliert Märkte, Informationsflüsse und staatliche Handlungsfähigkeit.»

Big Tech, Informationsmonopole und Geopolitik

Wenn KI-Systeme dem Gemeinwohl, der Demokratie und der Nachhaltigkeit dienen sollen – statt Märkte, Staaten und Informationsflüsse zu dominieren –, müssen wir diesen Machtfaktor gezielt analysieren und gegensteuern. Ein zentraler Punkt der Analyse: Die ohnehin seit Jahren kritisierte Macht von Big Tech verschärft sich im KI-Sektor aktuell drastisch (vgl. Böff 2026). Die Allianzen zwischen Amazon, Anthropic, Microsoft und Co. werden immer tiefer. Während diese Kooperationen für die Konzerne strategisch sinnvoll sein mögen, entstehen gesellschaftlich gefährliche Abhängigkeiten. Mögliche Folgen sind wirtschaftliche Instabilitäten und Informationsmonopole von globaler und geopolitischer Tragweite.

China als aufstrebende Tech-Macht

Beispielsweise erweitert China über eine digitale Seidenstrasse systematisch seinen globalen Einfluss (Dachwitz und Hilbig 2024). Ein Hauptaugenmerk liegt auf Afrika: Die dortige junge, digitalisi-

sierungsaffine Bevölkerung bietet Tech-Unternehmen enorme Chancen für künftige Absatzmärkte. Es steht allerdings weit mehr auf dem Spiel als ausschliesslich ökonomische Opportunitäten – Peking geht es vielmehr um eine neue digitale Weltordnung, wie Kritiker meinen (ebenda). Nachgewiesene Spionageaktivitäten im Zuge afrikanischer Infrastrukturprojekte sind dabei nur ein Nebenschauplatz. Im Zentrum steht das massive Macht- und Manipulationspotenzial, das durch den Zugriff auf Daten entsteht. Eine entscheidende Frage der Zukunft ist daher, wer die Kontrolle über diejenigen Daten erhält, die durch digitale Dienste und KI-Systeme generiert werden.

Der Kampf um Datensouveränität

Daher kämpfen zivilgesellschaftliche Organisationen um Datensouveränität. Sie wollen verhindern, dass Daten aus Ländern wie Bangladesch, Kenia oder Guatemala unkontrolliert auf den Servern der grossen Tech-Konzerne landen (ebenda). Auch Regierungen in Indien, Nigeria oder Südafrika setzen sich zunehmend für digitale Souveränität ein, zu der ganz zentral die Speicherung von Daten auf lokalen Servern zählt. Eine interessante Rolle nehmen in diesem Kontext die USA ein: Die Grossmacht agierte zunächst ganz im Sinne der heimischen Tech-Branche und stellte sich gegen jegliche Beschränkungen. Unter der zweiten Präsidentschaft von Joe Biden änderte sich diese Position jedoch – mit der offiziellen Begründung, dass innenpolitisch regulatorischer Spielraum benötigt werde (Congress.gov 2025). Gleichwohl spielte dabei auch die strategische Konkurrenz zu China eine entscheidende Rolle (Mok 2025).

Die widersprüchliche Politik der Europäischen Union

Interessanterweise unterstützte die Europäische Union (EU) das Engagement vieler Länder für Datensouveränität weitaus weniger, als zu erwarten gewesen wäre. Die Ursache dafür könnte Lobbyismus sein (Dachwitz und Hilbig 2024). Denn einen offensichtlichen Grund gibt es nicht, da in dem Staatenverbund keine Big-Techs und kaum KI-Unternehmen angesiedelt sind. Auch zu den eigenen Regulierungsmassnahmen

im Bereich der KI-Systeme scheint dies im Widerspruch zu stehen. Denn als wegweisend gilt hierbei etwa das europäische KI-Gesetz (EU AI Act), das 2024 in Kraft trat und dessen Durchsetzung und Umsetzung von einem eigens eingerichteten «Amt für künstliche Intelligenz» überwacht wird (EU-Kommission o.D.). Zentrale Elemente des EU AI Act umfassen Verbote bestimmter, etwa manipulativer, Anwendungen, strenge Vorgaben für Hochrisiko-KI, beispielsweise in den Bereichen Gesundheit und Bildung, sowie Transparenzpflichten und Prüf- oder Sicherheitsvorkehrungen (European Parliament 2023). Über die Transparenzforderungen des EU AI Act dürfte zudem auch das Problem der Deepfakes, KI-Systeme, die sich als Menschen ausgeben, adressiert werden. Ein Verbot, welches auch zunehmend in den USA gefordert wird (Gary 2026).

Europa als «führende Grossmacht» der Digitalregulierung

Die EU gilt über die Regulierung von KI-Systemen hinaus in ihrer Digitalpolitik als fortgeschritten. So schreiben Dachwitz und Hilbig (2024, S. 275): «In Sachen Digitalregulierung ist Europa die führende Grossmacht.» Keine Region der Welt habe grössere Anstrengungen unternommen, die unkontrollierte Macht der Digitalkonzerne einzuhegen und Bürgerinnen und Bürger vor deren Einfluss zu schützen (ebenda). Von diesem Privileg profitierten allerdings nur die Menschen im EU-Binnenraum, lautet die Kritik. Zudem gebe der Staatenbund strategischen Belangen Vorrang vor wertegestütztem Handeln – etwa mit Blick auf den Schutz Oppositioneller in repressiven Regimen. Ihr «Dritter Weg der Digitalisierung» (ebenda, S. 276) sei daher wenig altruistisch motiviert. Vielmehr habe auch die EU – wie China – die geopolitische Relevanz erkannt und entwickle sich selbst immer mehr zum «Akteur des digitalen Kolonialismus» (ebenda).

Märkte zwischen Euphorie, Angst und vorsichtigem Optimismus

Welche geopolitische, gesellschaftliche und ökonomische Relevanz KI-Systemen künftig tatsächlich zukommen wird, lässt sich allerdings schwer ab-

schätzen. Aus Investorensicht verweisen Milmo und Down (2026) darauf, dass grosse Technologiekonzerne zwar enorme Summen in KI-Infrastruktur investieren, zugleich aber Zweifel angebracht sind, ob sich diese Investitionen kurzfristig rentieren werden. Einige Analysen sehen Anzeichen für eine mögliche Überbewertung oder sogar eine Spekulationsblase rund um KI-Systeme. Während einige Fachleute glauben, dass die Auswirkungen von KI-Systemen aktuell überschätzt werden, meinen andere, die langfristigen Veränderungen würden noch unterschätzt. Die Stimmung schwankt zwischen Euphorie, Angst und vorsichtigem Optimismus.

Wichtige Rolle verantwortlich handelnder Investoren

Fest steht jedoch, dass KI-Systeme unsere Lebensrealitäten weiter verändern werden. Wir müssen daher Wege finden, wie KI-Systeme nachhaltig werden und in den Dienst einer nachhaltigen Entwicklung gestellt werden können. Verantwortlich auch und gerade im Sinne der treuhänderischen Pflichten handelnde Investoren können hierzu einen wichtigen Beitrag leisten, wenn sie die relevanten ökologischen, sozialen und Governance-bezogenen Aspekte in ihre Analysen einbeziehen. Und zwar sollten hierbei sowohl die regulatorisch-politische als auch die institutionelle und individuelle Ebene Berücksichtigung finden. Gerade letztere kann sich als besondere Herausforderung, aber auch als Chance erweisen.

Von Personal, Talenten und «Figuren»

Laut dem Wissenschaftler Xiong (2026) lassen sich vor dem Hintergrund von KI-Systemen drei Schichten unterscheiden: Personal, Talente und «Figuren»: Das Personal verrichtet routinemässige und vorhersehbare Arbeiten, weswegen es künftig durch KI-Systeme ersetzt wird. Talente besitzen eine tiefe Expertise, verbunden mit der Bereitschaft zum lebenslangen Lernen. Sie werden laut dem Wissenschaftler aus Hongkong aber dann verdrängt werden, wenn sie ihre eigene Produktivität nicht durch KI-Systeme steigern. Besonders aber kommt es laut Xiong auf die «Fi-

guren» an, die über blosse fachliche Meisterschaft hinaus tiefgreifende strategische Weitsicht und Führungskraft in sich vereinen und daher Aufgaben wahrnehmen, die Algorithmen nicht abdecken können.

Vom Mut, alte Denkmuster

zu zerschlagen, zur digitalen Utopie Zweifellos besteht somit die Gefahr, dass sich durch KI-Systeme bestehende gesellschaftliche Schichtunterschiede verstärken oder neu entstehen. Und es ist eine Aufgabe, einem solchen Trend entgegenzutreten – etwa durch geeignete Bildungsmassnahmen. Zugleich ist mit KI-Systemen aber auch das grosse Potenzial verbunden, dass sie unsere originär menschlichen Fähigkeiten und damit uns alle indirekt stärken. Xiong formuliert dies so: Wenn die Routine-Aufgaben durch KI-Systeme erledigt werden, verbleiben zutiefst menschliche Eigenschaften wie die Neugier auf das Unbekannte, emotionale und moralische Tiefe sowie der Mut, alte Denkmuster zu zerschlagen. Und vielleicht benötigen wir gerade Letzteres auch in Bezug auf KI-Systeme und die Digitalisierung. Wenn dies so ist, sollten wir uns für die positiven Visionen wie derjenigen von Renata Ávila Pinto, Anwältin für Menschenrechte und Technologie, öffnen, die eine digitale «Utopie wider den digitalen Kolonialismus» formuliert (Dachwitz und Hilbig 2024).

*luisa.lange@ethius-invest.ch
www.ethius-invest.ch*

Literaturverzeichnis zu «Ethik, Governance und Regulierung von KI-Systemen: Eine ungeklärte Machtfrage»

- Böff, Melanie (2026): Boom bei Künstlicher Intelligenz. Die riskanten Allianzen der Big-Tech-Konzerne. Erschienen am 23. April 2026 auf tagesschau.de. <https://www.tagesschau.de/wirtschaft/digitales/allianz-der-techkonzerne-100.html>
- Chee, Foo Yun und Hummel, Tassilo (2024): Europe Sets Benchmark for Rest of the World with Landmark AI laws. <https://www.reuters.com/world/europe/eu-countries-back-landmark-artificial-intelligence-rules-2024-05-21/>
- Congress.gov (2025): Digital Trade and Data Policy: Key Issues Facing Congress. <https://www.congress.gov/crs-product/IF12347>
- Dachwitz, Ingo und Hilbig, Sven (2024): Digitaler Kolonialismus. Wie Tech-Konzerne und Grossmächte die Welt unter sich aufteilen. C.H.Beck, München.
- EU-Kommission (o.D.): Governance und Durchsetzung des KI-Gesetzes. <https://digital-strategy.ec.europa.eu/de/policies/ai-act-governance-and-enforcement>
- European Parliament (2025): EU AI Act: First Regulation on Artificial Intelligence. <https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence>
- Fratzscher, Marcel (2026): Einfluss von KI: Künstliche Intelligenz bedroht die Demokratie, aber anders als gedacht. Erschienen in Die Zeit am 24. April 2026. <https://www.zeit.de/wirtschaft/2026-04/einfluss-ki-macht-verteilung-oekonomie>
- IÖW (2021): Nachhaltigkeitskriterien für Künstliche Intelligenz. Entwicklung eines Kriterien- und Indikatorensets für die Nachhaltigkeitsbewertung von KI-Systemen entlang des Lebenszyklus. https://www.ioew.de/fileadmin/user_upload/Bilder_und_Downloaddateien/Publikationen/2021/IOEW_SR_220_Nachhaltigkeitskriterien_fuer_Kuenstliche_Intelligenz.pdf
- Marcus, Gary (2026): We Urgently Need a Federal Law Forbidding AI from Impersonating Humans. <https://garymarcus.substack.com/p/we-urgently-need-a-federal-law-forbidding>
- Milmo, Dan and Down, Aisha (2026): AI is Indeed Coming – But there is also Evidence to Allay Investor Fears. Erschienen in The Guardian vom 13.02.2026. <https://www.theguardian.com/technology/2026/feb/13/ai-effects-on-business-industry-evidence-to-allay-investor-fears>
- Mok, Charles (2025): China Picked Up the Ball on Data Trade Where the US Dropped It in 2024. <https://www.techpolicy.press/china-picked-up-the-ball-on-data-trade-where-the-us-dropped-it-in-2024/>
- Rachovitsa, Adamantia et al. (2022): The Human Rights Implications of the Use of AI in the Digital Welfare State: Lessons Learned from the Dutch SyRI Case. <https://academic.oup.com/hrlr/article/22/2/ngac010/6568079>
- Singh, Kanishka (2024): U.S. Looking at Report that Israel used AI to Identify Bombing Targets in Gaza. Erschienen in Reuters am 5. April 2024. <https://www.reuters.com/world/middle-east/us-looking-report-that-israel-used-ai-identify-bombing-targets-gaza-2024-04-04/>
- Xiong, Hui (2026): Mensch gegen Maschine: Silizium-Kolosse gegen Kohlenstoff-Leben. Erschienen in der Frankfurter Allgemeinen Zeitung vom 04.05.2026. <https://www.faz.net/aktuell/wirtschaft/unternehmen/mensch-gegen-maschine-wo-ki-an-ihre-grenzen-stoesst-accg-200777081.html>