



Europäische KI-Regulierung: Auf der Suche nach nachhaltigen und inklusiven Ansätzen für gemeinsame IT-Welten.

Marco Wedel (TU Berlin), Alexandra Wudel (FemAI) und Antonios Hazim (nexus Institut)
Forum Privatheit 2023: Data Sharing – Datenkapitalismus by Default? | 05. Oktober 2023



Dr. Marco Wedel

TU Berlin, Chair for Work,
Technology and Participation,
Berlin, Germany



Alexandra Wudel

FemAI – Center for Feminist
Artificial Intelligence GmbH



Antonios Hazim

nexus - Institut für
Kooperationsmanagement u.
interdisziplinäre Forschung

EU KI-Regulierung: Auf der Suche nach nachhaltigen und inklusiven Ansätzen für gemeinsame IT-Welten.



Agenda:

1. Grund für und Einsicht in Regulierungsnotwendigkeit (Warum?)
2. Ansatz der Regulierungsbemühungen und Willensbildungsprozess (Was?)
3. Umsetzung der Handlungsabsichten und -ansprüchen (Was/Wie?)
4. (Zusätzliche) Impulse für Nachhaltige und inklusive Handlungsansätze (Was/Wie?)

Fokus:

Nach nachhaltigen und inklusiven Ansätzen für gemeinsame IT-Welten!

Warum? (Agenda Setting)

European Economic and Social Committee (2016/2017)



Own-initiative opinion:

The impact of AI is of a cross-border nature and therefore supra-national policy frameworks will also need to be established. The EESC recommends that the EU take the **lead globally in establishing clear global policy frameworks for AI, in line with European values and fundamental rights.**

The EESC currently **identifies 11 areas where AI poses societal challenges**: ethics; safety; privacy; transparency and accountability; work; education and skills; (in)equality and inclusiveness; law and regulations; governance and democracy; warfare; superintelligence.

The EESC calls for a **human-in-command approach to AI**, including the precondition that the development of AI be responsible, safe and useful, where machines remain machines and people retain control over these machines at all times.

The EESC calls for a **code of ethics** (...) development of a **standardisation system** (...) **European AI infrastructure.**

Dok: European Economic and Social Committee. Artificial intelligence – The consequences of artificial intelligence on the (digital) single market, production, consumption, employment and society (own-initiative opinion).

Warum? (Agenda Setting)

European Parliament (2016/2017)



Resolution:

From Mary Shelley's **Frankenstein's Monster** to the classical myth of **Pygmalion**, through the story of Prague's **Golem** to the **robot** of Karel Čapek (...)

Considers it essential, in the development of robotics and AI, to **guarantee that humans have control over intelligent machines at all times**(...)

Whereas now that humankind stands on the threshold of an era when ever more sophisticated robots, bots, androids and other manifestations of artificial intelligence ('AI') seem to be poised to unleash a **new industrial revolution**, which is likely to leave no stratum of society untouched, it is vitally important for the legislature to **consider its legal and ethical implications** and effects, without stifling innovation(...)

Considers that the **existing Union legal framework should be updated** and complemented, where appropriate, by guiding ethical principles in line with the complexity of robotics and its many social, medical and bioethical implications(...)

Dok: European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL))

Warum? (Agenda Setting)

Council and Commission (2016/2017)



Council:

(...)a **sense of urgency** to address emerging trends: this includes issues such as artificial intelligence and blockchain technologies, while at the same time ensuring a high level of **data protection, digital rights and ethical standards**. The European Council invites the Commission to put forward a European approach to artificial intelligence by early 2018 (...) to **explore new markets through risk-based radical innovations and to reaffirm the leading role of its industry.**”

Commission (Juncker):

It (EU) has a comprehensive legal framework which protects consumers while promoting innovation and it is making progress in creating a Digital Single Market. **The main ingredients are there for the EU to become a leader in the AI revolution**, in its own way and based on its values.

“(t)he Commission appointed an **independent AI high-level expert group** with the task of developing draft AI ethics guidelines(...)”



Warum? (Agenda Setting)

High-Level Expert Group on AI (2019)



Ansatz:

Vertrauenswürdige, auf den Menschen ausgerichtete KI als ethisches *und* demokratisches Prinzip und innovative, ethische KI-Technologien als Wettbewerbsvorteil.

Eine vertrauenswürdige KI zeichnet sich durch drei Komponenten aus:

- Sie sollte **rechtmäßig** sein und somit geltendes Recht und alle gesetzlichen Bestimmungen einhalten;
- Sie sollte **ethisch** sein und somit die Einhaltung ethischer Grundsätze und Werte garantieren;
- Sie sollte **robust** sein, und zwar sowohl **in technischer als auch in sozialer Hinsicht**, da KI-Systeme möglicherweise unbeabsichtigten Schaden verursachen, selbst wenn ihnen gute Absichten zugrunde liegen.

Jede dieser drei Komponenten ist notwendig, jedoch alleine nicht ausreichend, um das Ziel einer vertrauenswürdigen KI zu erreichen.

Dok: Unabhängige Hochrangige Expertengruppe für Künstliche Intelligenz. Eingesetzt von der Europäischen Kommission im Juni 2018. Ethikleitlinien für eine vertrauenswürdige KI. 2019

Warum? (Agenda Setting)

High-Level Expert Group on AI (2019)

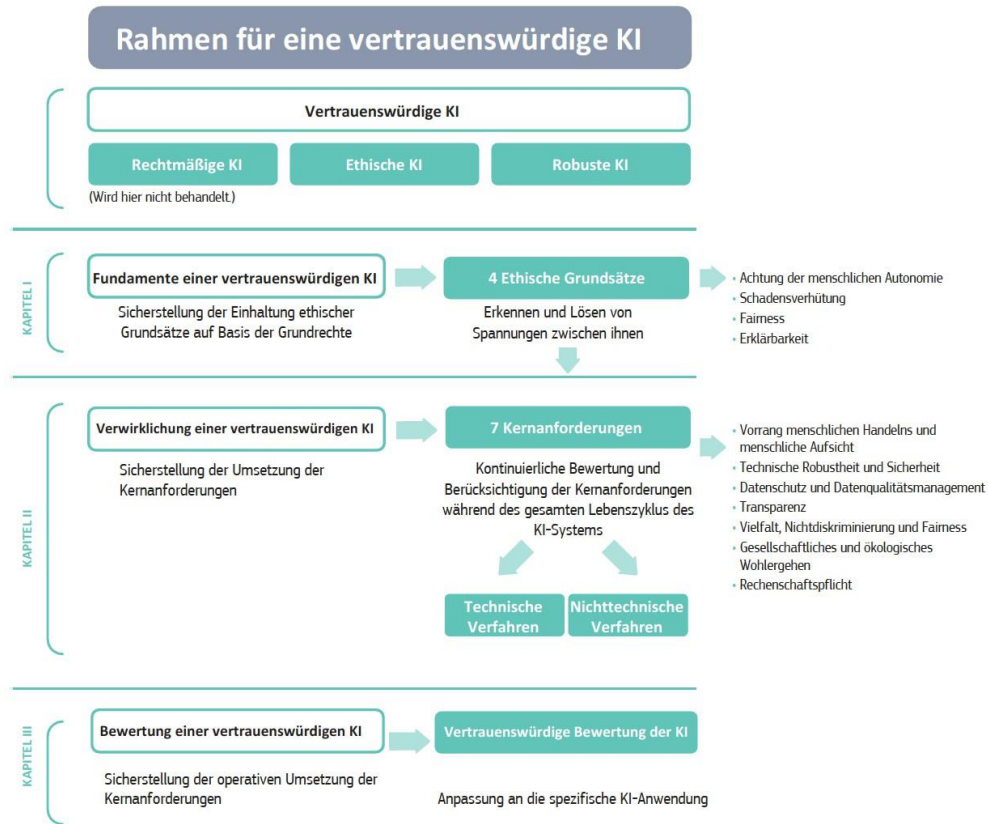


Abbildung 1: Die Leitlinien als Rahmen für eine vertrauenswürdige KI

Dok: Unabhängige Hochrangige Expertengruppe für Künstliche Intelligenz. Eingesetzt von der Europäischen Kommission, Juli 2018. Ethical Guidelines for trustworthy AI. 2019

4 Ethische Grundsätze:

- Achtung der menschlichen Autonomie
- Schadensverhütung
- Fairness
- Erklärbarkeit

7 Kernanforderungen:

- Vorrang menschlichen Handelns und menschliche Aufsicht
- Technische Robustheit und Sicherheit
- Datenschutz und Datenqualitätsmanagement
- Transparenz

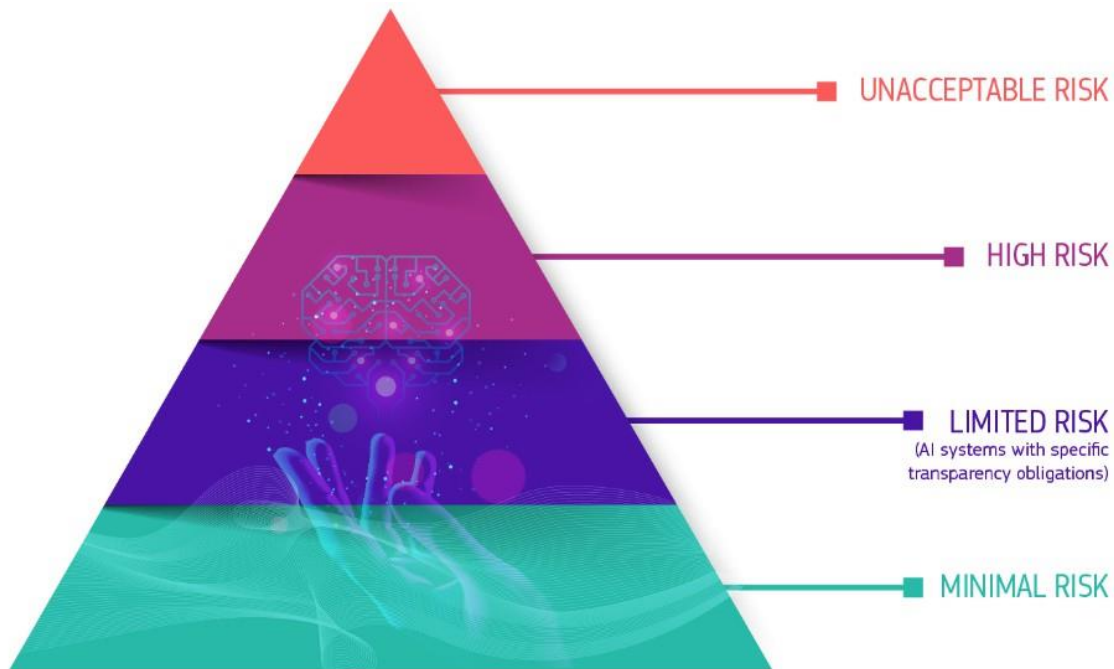
Vielfalt, Nichtdiskriminierung und Fairness

- Gesellschaftliches und ökologisches Wohlergehen
- Rechenschaftspflicht



Was? (Gesetzgebungsverfahren)

Commission Draft EU AI Act (2021)



- Alles, was als eindeutige Bedrohung für EU-Bürger angesehen wird (**physischer, psychischer Schaden**) und dem „**human-in-command**“ Ansatz widerspricht.
- Acht definierte Hochrisiko-Bereiche. Darunter **Beschäftigung, Personalmanagement und Zugang zu selbstständiger Tätigkeit** (z.B. Software zur Auswertung von Lebensläufen).
- KI-Systeme, die für die **Interaktion mit natürlichen Personen** bestimmt sind. Transparenzverpflichtungen, damit Nutzer:innen entscheiden können, ob sie die Anwendung weiter nutzen oder nicht.
- Etwa kostenlose Nutzung von Anwendungen wie KI-gestützten Videospielen oder Spamfiltern.

Dok: Unabhängige Hochrangige Expertengruppe für Künstliche Intelligenz. Eingesetzt von der Europäischen Kommission im Juni 2018. Ethikleitlinien für eine vertrauenswürdige KI. 2019

Was? (Gesetzgebungsverfahren)

Drei Institutionen und jede Menge Lobbyismus (seit 2021)



Fragen, die im Gesetzgebungsverfahren kontrovers diskutiert (und hier nicht im Fokus stehen):

- Was ist KI?
- Was sind verbotene Anwendungsbereiche?
- Welche Anwendungsbereiche sind mit einem hohen Risiko behaftet?
- Wie lässt sich Generative AI (GenAI) und General Purpose AI (GPAI) verorten?
- Was ist mit AI Literacy?
- Wie umgehen mit biometrischer Fernidentifizierung?
- Welche Maßnahmen sind zu ergreifen?
 - Datenqualitätsmanagement
 - Transparenz
 - Erklärbarkeit
- Welche Maßnahmen sind verpflichtend, welche freiwillig zu ergreifen?
- Wer trägt die Verantwortung?



Was? (Gesetzgebungsverfahren)

Nachhaltige und inklusive Gestaltung von KI-Systemen?



Der Ansatz der drei Institutionen zu inklusiven und nachhaltigen Gestaltungsansätzen:

4 Ethische Grundsätze:

- Achtung der menschlichen Autonomie
- Schadensverhütung
- Fairness
- Erklärbarkeit

7 Kernforderungen:

- Vorrang menschlichen Handelns und menschliche Aufsicht
- Technische Robustheit und Sicherheit
- Datenschutz und Datenqualitätsmanagement
- Transparenz
- Vielfalt, Nichtdiskriminierung und Fairness
- Gesellschaftliches und ökologisches Wohlergehen
- Rechenschaftspflicht

Schaut man sich die Verordnungsentwürfe der EU-Kommission (2021), des Rates der EU (2022) und des EU-Parlaments (2023) an, **findet sich der zuvor beschriebenen Anspruch**, hier insbesondere die von der Hochrangigen Expertengruppe für vertrauenswürdige KI definierten ethischen Ansprüche und Kernforderungen für eine vertrauenswürdige KI (HEKI 2019), **mehr oder weniger implizit in den Gesetzestexten wieder**. Im Entwurf des EU-Parlaments werden diese sogar explizit referenziert – *Art. 4a. General principles applicable to all AI systems* (EP 2023, S. 142).

Was? (Gesetzgebungsverfahren)

ABER nachhaltige und inklusive Gestaltung freiwillig!



Entwurf Kommission (Kap. 2, Art. 69) April 2021:

1. Die Kommission und die Mitgliedstaaten fördern und erleichtern die Aufstellung von **Verhaltenskodizes, mit denen erreicht werden soll**, dass die in Titel III Kapitel 2 genannten **Anforderungen auf KI-Systeme** Anwendung finden, **die kein hohes Risiko bergen** (...).
2. Die Kommission und der Ausschuss fördern und erleichtern die Aufstellung von Verhaltenskodizes, mit denen erreicht werden soll, dass KI-Systeme **freiwillig weitere Anforderungen erfüllen**, die sich beispielsweise auf die ökologische Nachhaltigkeit, die barrierefreie Zugänglichkeit für Personen mit Behinderungen, die Beteiligung von Interessenträgern an der Konzeption und Entwicklung von KI-Systemen und die Vielfalt der Entwicklungsteams beziehen, wobei die Erreichung dieser Ziele anhand klarer Vorgaben und wesentlicher Leistungsindikatoren gemessen wird.
3. Verhaltenskodizes können von einzelnen **KI-System-Anbietern oder von Interessenvertretungen dieser Anbieter oder von beiden aufgestellt werden**, auch unter Einbeziehung von Nutzern und Interessenträgern sowie deren Interessenvertretungen. Verhaltenskodizes können sich auf mehrere KI-Systeme erstrecken, um ähnlichen Zweckbestimmungen der jeweiligen Systeme Rechnung zu tragen.



Was? (Gesetzgebungsverfahren)

ABER nachhaltige und inklusive Gestaltung freiwillig!



Entwurf Rat der Europäischen Union (Art. 69) November 2022:

1. The Commission, and the Member States shall facilitate the drawing up of codes of conduct intended to encourage the voluntary application to AI systems other than high-risk AI systems of one or more of the requirements set out in Title III, Chapter 2 of this Regulation to the best extent possible, taking into account the available, technical solutions allowing for the application of such requirements.
2. The Commission and the Member States shall facilitate the drawing up of codes of conduct intended to encourage the voluntary application to all AI systems of specific requirements related, for example, to environmental sustainability, including as regards energy-efficient programming, accessibility for persons with a disability, stakeholders participation in the design and development of the AI systems and diversity of development teams on the basis of clear objectives and key performance indicators to measure the achievement of those objectives. The Commission and the Member States shall also facilitate, where appropriate, the drawing of **codes of conduct** applicable on a voluntary basis **with regard to users' obligations** in relation to AI systems.



Was? (Gesetzgebungsverfahren)

ABER nachhaltige und inklusive Gestaltung freiwillig!



Entwurf Europäisches Parlament (Art. 69) Mai/Juni 2023:

Codes of conduct intended to foster the voluntary compliance with the principles underpinning trustworthy AI systems, shall, in particular:

- a) **aim for a sufficient level of AI literacy** among their staff and other persons dealing with the operation and use of AI systems in order to observe such principles;
- b) **assess to what extent their AI systems may affect vulnerable persons or groups of persons**, including children, the elderly, migrants and persons with disabilities or whether measures could be put in place in order to increase accessibility, or otherwise support such persons or groups of persons;
- c) **consider** the way in which the use of their **AI systems may have an impact or can increase diversity, gender balance and equality**;
- d) **have regard to whether their AI systems** can be used in a way that, directly or indirectly, may residually or significantly **reinforce existing biases or inequalities**;
- e) **reflect on the need and relevance of having in place diverse development teams** in view of securing an inclusive design of their systems;
- f) **give careful consideration to whether their systems can have a negative societal impact**, notably concerning political institutions and democratic processes;
- g) **evaluate how AI systems can contribute to environmental sustainability** and in particular to the Union's commitments under the European Green Deal and the European Declaration on Digital Rights and Principles.

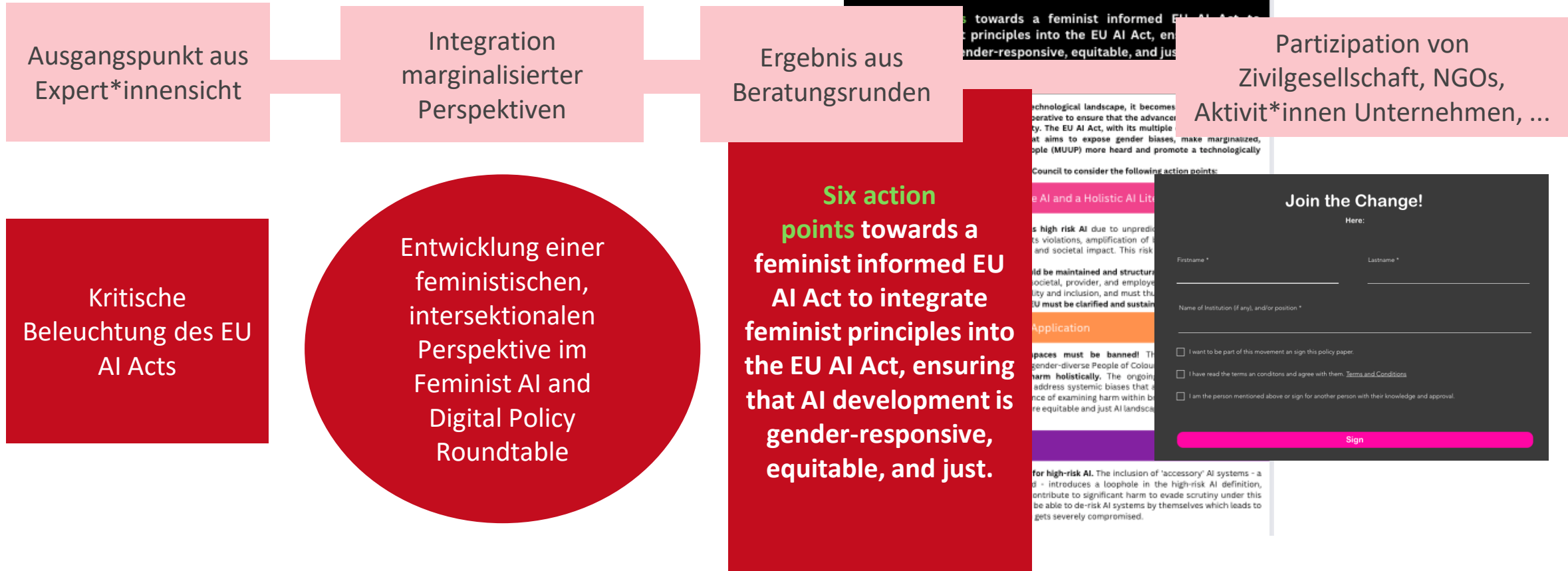
Unser Fazit:

**Nachhaltige und inklusive Ansätze sind
in der europäischen KI-Regulierung ein blinder Fleck.**

(vgl. hierzu auch Hacker 2023 in Bezug auf AI Liability Directive)

Impulse für nachhaltige und inklusive Handlungsansätze

– der Prozess



Impulse für nachhaltige und inklusive Handlungsansätze – die Inhalte



A feminist vision for the EU AI Act  FemAI - Center for Feminist Artificial Intelligence

Six action points towards a feminist informed EU AI Act to integrate feminist principles into the EU AI Act, ensuring that AI development is gender-responsive, equitable, and just.

As society navigates this transformative technological landscape, it becomes increasingly evident that the application of a feminist lens is imperative to ensure that the advancements in AI align with principles of equality, justice, and inclusivity. The EU AI Act, with its multiple dimensions, provides fertile ground for a feminist analysis that aims to expose gender biases, make marginalized, underrepresented, and underprivileged people (MUUP) more heard and promote a technologically empowered world for all.

Therefore, we ask the EU Parliament and EU Council to consider the following action points:

- 1. Classification of General Purpose AI and a Holistic AI Literacy**
 - **General Purpose AI needs to be scored as high risk AI** due to unpredictable and uncontrollable outcomes which may result in human rights violations, amplification of biases and discrimination, complex interactions as well as scalability and societal impact. This risk is significantly higher for MUUP.
 - **The current holistic view of AI literacy should be maintained and structurally funded by the EU.** The Parliament's emphasis on AI literacy from societal, provider, and employer perspectives aligns well with feminist aims concerning intersectionality and inclusion, and must thus be kept up in the future EU AI Act. However, **the specific role of the EU must be clarified and sustained with financial impact.**
- 2. Identifying Prohibited Areas of Application**
 - **Facial recognition software in public spaces must be banned!** The risk of intersectional discrimination, in particular for Women and gender-diverse People of Colour is unacceptable.
 - **The need to understand and measure harm holistically.** The ongoing standardization of risk assessment must be vigilantly observed to address systemic biases that amplify harm for MUUP. A feminist viewpoint underscores the importance of examining harm within broader social, cultural, and historical contexts, thereby striving for a more equitable and just AI landscape.
- 3. No Loopholes for High-Risk AI**
 - **Identify, and pre-empt, possible loopholes for high-risk AI.** The inclusion of 'accessory' AI systems - a concept that remains ambiguously defined - introduces a loophole in the high-risk AI definition, potentially permitting systems that might contribute to significant harm to evade scrutiny under this classification. For instance, providers would be able to de-risk AI systems by themselves which leads to less rules to follow. The risk-based approach gets severely compromised.

Our Policy Paper contains six action points towards a feminist informed EU AI Act to integrate feminist principles into the EU AI Act, ensuring that AI development is gender-responsive, equitable, and just.

1. Classification of General Purpose AI and a Holistic AI Literacy

2. Identifying Prohibited Areas of Application

3. No Loopholes for High-Risk AI

4. Inclusive Data Sets setting high standards while lacking practicability

5. Ensuring transparency and accountability through the AI Office

6. Rejecting Voluntary Aspects: A Feminist Critique



Impulse für nachhaltige und inklusive Handlungsansätze – die Dringlichkeit



Demonstration von ungleichverteilten Machtverhältnissen im patriarchalen System

Ausblick

Partizipation von Zivilgesellschaft, Unternehmen, ...

Kompromissfindung: für die weniger konfliktreichen Fragen anlässlich des politischen Trilogs am 2./3. Oktober 2023

Zu den komplexeren Themen sollen anlässlich des politischen Trilogs am 25. Oktober 2023

Lobbypower von US-Tech-Konzernen

Schnelle Trilogverhandlungen lassen weniger Reaktionsraum für die Zivilgesellschaft

Und die Standardisierung ?

A feminist vision for the EU

Six action points towards a feminist vision for the EU. Integrate feminist principles into AI development is gender-responsive, equitable, and just.

As society navigates this transformative technological landscape, it becomes increasingly evident that the application of a feminist lens is imperative to ensure that the advancements in AI align with principles of equality, justice, and inclusivity. The EU AI Act, with its multiple dimensions, provides fertile ground for a feminist analysis that aims to expose gender biases, make marginalized, underrepresented, and underprivileged people (MUUP) more heard and promote a technologically empowered world for all. Therefore, we ask the EU Parliament and EU Council to consider the following action points:

1. Classification of General Purpose AI and a Holistic AI Literacy

- **General Purpose AI needs to be scored as high risk AI** due to unpredictable and uncontrollable outcomes which may result in human rights violations, amplification of biases and discrimination, complex interactions as well as scalability and societal impact. This risk is significantly higher for MUUP.
- **The current holistic view of AI literacy should be maintained and structurally funded by the EU.** The Parliament's emphasis on AI literacy from societal, provider, and employer perspectives aligns well with feminist aims concerning intersectionality and inclusion, and must thus be kept up in the future EU AI Act. However, **the specific role of the EU must be clarified and sustained with financial impact.**

2. Identifying Prohibited Areas of Application

- **Facial recognition software in public spaces must be banned!** The risk of intersectional discrimination, in particular for Women and gender-diverse People of Colour is unacceptable.
- **The need to understand and measure harm holistically.** The ongoing standardization of risk assessment must be vigilantly observed to address systemic biases that amplify harm for MUUP. A feminist viewpoint underscores the importance of examining harm within broader social, cultural, and historical contexts, thereby striving for a more equitable and just AI landscape.

3. No Loopholes for High-Risk AI

- **Identify, and pre-empt, possible loopholes for high-risk AI.** The inclusion of 'accessory' AI systems - a concept that remains ambiguously defined - introduces a loophole in the high-risk AI definition, potentially permitting systems that might contribute to significant harm to evade scrutiny under this classification. For instance, providers would be able to de-risk AI systems by themselves which leads to less rules to follow. The risk-based approach gets severely compromised.

KI-Regulierung: Auf der Suche nach nachhaltigen und inklusiven Ansätzen für gemeinsame IT-Welten.



Danke für Ihre Aufmerksamkeit!



*Unterschreibe das Policy Paper für eine
feministische Vision von AI*

Literatur



Council of the European Union (2022): Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts - General, data.consilium.europa.eu/doc/document/ST-14954-2022-INIT/en/pdf (Abfrage 28.05.2023)

European Parliament (2020): Legislative Entschließung des Europäischen Parlaments vom 20. Oktober 2020 mit Empfehlungen an die Kommission zu dem Rahmen für die ethischen Aspekte von künstlicher Intelligenz, Robotik und damit zusammenhängenden Technologie. europarl.europa.eu/doceo/document/TA-9-2020-0275_DE.html (Abfrage 28.05.2023)

European Parliament (2023): DRAFT Compromise Amendments on the Draft Report Proposal for a regulation of the European Parliament and of the Council on harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts (COM(2021)0206 – C9 0146/2021 – 2021/0106(COD)). www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/CJ40/DV/2023/05-11/ConsolidatedCA_IMCOLIBE_AI_ACT_EN.pdf (Abfrage 28.05.2023)

Europäische Kommission (2020): Weissbuch. Zur Künstlichen Intelligenz – ein europäisches Konzept für Exzellenz und Vertrauen, ec.europa.eu/info/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_de (Abfrage 28.05.2023)

Europäische Kommission (2021): Vorschlag für eine Verordnung des Europäischen Parlaments und des Rates zur Festlegung harmonisierter Vorschriften für Künstliche Intelligenz (Gesetz zur Künstlichen Intelligenz) und zur Änderung bestimmter Rechtsakte der Union {SEC(2021) 167 final} - {SWD(2021) 84 final} - {SWD(2021) 85 final}, eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206 (Abfrage 15.05.2023)

Europäischer Wirtschafts- und Sozialausschuss (2016): Künstliche Intelligenz – die Auswirkungen der künstlichen Intelligenz auf den (digitalen) Binnenmarkt sowie Produktion, Verbrauch, Beschäftigung und Gesellschaft (Initiativstellungnahme), eesc.europa.eu/en/our-work/opinions-information-reports/opinions/artificial-intelligence-consequences-artificial-intelligence-digital-single-market-production-consumption-employment-and (Abfrage 15.05.2023)

Hacker, Philipp, The European AI Liability Directives – Critique of a Half-Hearted Approach and Lessons for the Future (November 25, 2022). Available at SSRN: <https://ssrn.com/abstract=4279796> or <http://dx.doi.org/10.2139/ssrn.4279796>

Hochrangige Expertengruppe (2019): Unabhängige Hochrangige Expertengruppe für Künstliche Intelligenz. Eingesetzt durch die Europäische Kommission im Juni 2018. Ethik-Leitlinien für eine vertrauenswürdige KI, op.europa.eu/de/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1 (Abfrage 17.05.2023)