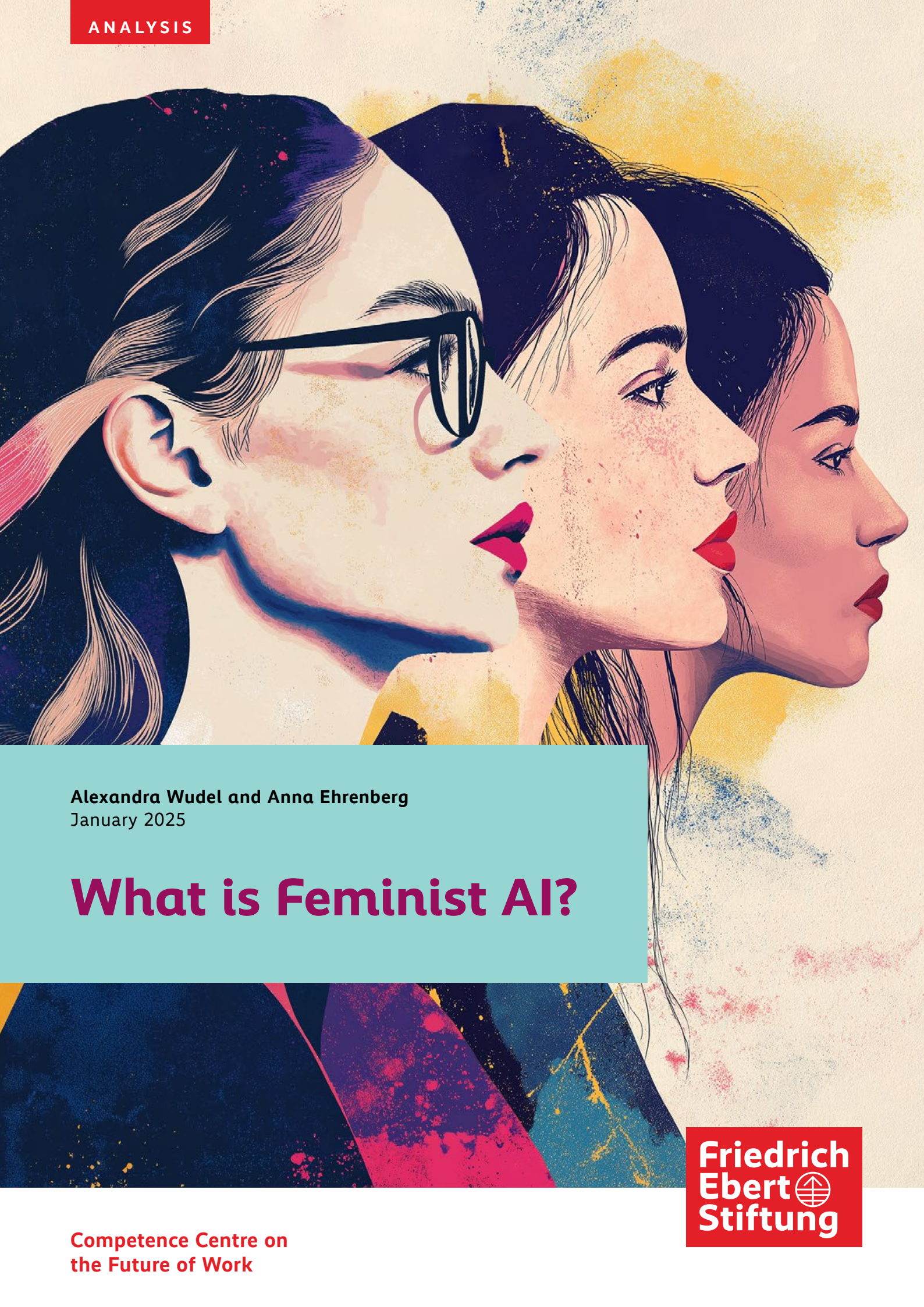




Alexandra Wudel and Anna Ehrenberg
January 2025

What is Feminist AI?

Imprint

Published by

Friedrich-Ebert-Stiftung e.V.
Godesberger Allee 149
53175 Bonn, Germany
info@fes.de

Issuing Department

Competence Centre on the Future of Work
Cours Saint Michel 30a, 1040 Brussels, Belgium

For more information about the Competence Centre
on the Future of Work, please consult:
<https://www.futureofwork.fes.de>

Responsibility for Content and Editing

Dr. Inga Sabanova
inga.sabanova@fes.de

Design/Layout

pertext | corporate publishing
www.pertext.de

The views expressed in this publication are not necessarily those of
the Friedrich-Ebert-Stiftung (FES). Commercial use of media published
by the FES is not permitted without the written consent of the FES.
Publications by the FES may not be used for electioneering purposes.

January 2025
© Friedrich-Ebert-Stiftung e.V.

ISBN 978-3-98628-688-0

Further publications of the Friedrich-Ebert-Stiftung can be found here:
[↗ www.fes.de/publikationen](http://www.fes.de/publikationen)

Alexandra Wudel and Anna Ehrenberg
January 2025

What is Feminist AI?

Content

Summary	3
Introduction	3
Background	4
Intersection between feminism and AI	4
How to bridge the gap between theory and practice	5
Ensuring AI systems seek to reduce power inequalities through feminist methodology	5
Ensuring AI systems meet societal demands and demonstrate implementation potential	5
Ensuring AI systems are continuously developed to reflect changing societal influences	5
Feminist AI in practice	6
Conclusion	6
References	7

Summary

This paper explores Feminist Artificial Intelligence (FAI), a framework leveraging intersectional feminism to address biases and inequities in AI systems. FAI emphasises interdisciplinary collaboration, systemic power analysis and iterative theory-practice loops. By embedding feminist values – equity, freedom and justice – FAI seeks to transform AI development, ensuring inclusivity and social sustainability. Practical applications include initiatives such as FemAI’s advocacy for feminist perspectives in the EU AI Act and the MIRA diagnostic platform, which aligns AI tools with social justice goals. FAI marks a critical departure from traditional AI by tackling structural inequalities and promoting accountable and equitable AI solutions.

Introduction

The intersection of feminism and artificial intelligence (AI) has emerged as a critical area for ensuring social equity within technological advancements. Feminist Artificial Intelligence (FAI) explores how feminist values, such as freedom, equity and justice, can influence AI development, challenging the male-dominated, Western-centric biases that currently pervade AI systems. This paper investigates FAI’s core principles, its role in disrupting power structures, and how it compares to other frameworks such as environmental, social and governance (ESG) criteria.

In this paper, we share FAI’s state of the art from both a theoretical and practical viewpoint to stay ahead of the fast-paced AI developments.

Background

Intersectional feminism serves as one of the most effective research methods to unlock the black box of AI systems operating in a patriarchal system. In this system, power and authority are primarily held by men, often resulting in the subordination of women and marginalised groups.

Although there are multiple definitions of feminism, Ferguson (2017) identifies a set of values shared amongst them which can be categorised into three pillars.

- First, feminist theory acknowledges the world as a complex subject, as opposed to simplifying it through dualistic thinking of “either/or”.
- Second, feminist thinking is compared to fluid processes rather than static entities, acknowledging their dynamics and change.
- Third, feminism is concerned with understanding the current world (intellectual subject) as well as changing it (a political undertaking) through equity, freedom and justice movements.

Ferguson (2017) formulates the common goal of challenging current power relations, imagining better worlds and striving to achieve them. The main tools to pursue this goal are intersectionality, interdisciplinarity and the execution of theory-practice feedback loops.

Intersectional feminism refers to a framework that acknowledges the oppression of people shaped not only by their gender, but by other social categories such as ethnicity, class, sexuality, ability, age, religion or geography, which intersect (Hill Collins, 2010). In a globalised, transnational world, people can belong to multiple groups simultaneously (Purkayastha, 2012).

Theory-practice feedback loops are used to emphasise how both theory and practice constantly influence each other through feedback loops (Ferguson, 2017). Taking the intellectual subject and the political undertaking into account, FAI is motivated by envisioning a better world beyond “one-size-fits-all” solutions which typically miss cultural nuances (Arora and Chowdhury, 2021). This could be addressed through intersectional research and by setting up interdisciplinary teams.

Intersection between feminism and AI

The existence of the concept FAI implies that the current landscape of AI systems lacks alignment with intersectional feminist values. An AI system, as defined in the EU AI Act, refers to “a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments” (European Parliament and Council of the European Union, 2024a). It transfers societal structures from the analogue to the digital world and drives both risks and opportunities by influencing society with its outputs (Gillam and Jain, 2024; Toupin, 2024).

The present understanding of intelligence in the field of AI is described as mostly rational and masculinist (Atanasoski and Vora, 2019: 195). Additionally, the set of values that shape the design of AI consists of those based on Western cultures (Arora and Chowdhury, 2021: 5). As a result, AI currently leads to a digitised society embedded into a social system with power inequalities. Subsequently, AI systems may reinforce structural discrimination of marginalised, underrepresented and underprivileged people (MUUP) (Gengler et al., 2023; Toupin, 2024).

This raises the question of how FAI can help shift societal structures from patriarchal foundations toward more equitable futures, and to what extent the concepts of feminism and AI may remain misaligned.

How to bridge the gap between theory and practice

AI can be shaped by the idea of societal and environmental sustainability, creating a world where patriarchal structures are reduced, where there is equal pay for equal work, where more people have access to healthcare, and where nature is thriving – a peaceful world with minimal discrimination. In order to achieve this, three aspects must be considered.

Ensuring AI systems seek to reduce power inequalities through feminist methodology

This aspect emphasises the application of an **intersectional feminist methodology** in the field of AI. This involves creating interdisciplinary and diverse teams and fostering iterative theory-practice loops to ensure that theoretical principles translate effectively into real-world applications. Importantly, FAI extends beyond the direct development of AI systems; it involves engaging multiple stakeholders (Gillam and Jain, 2024), such as policymakers and individual users, to advocate for systemic change.

One practical tool for bridging the gap between principles and practice is the Feminist Design Tool by the Feminist Internet and Young (Feminist Internet and Young, 2020). This tool operationalises feminist principles by translating them into actionable questions about topics such as stakeholders, purpose and context during the design process.

Through this methodology, power inequalities are identified and addressed with targeted measures to create more equitable AI systems with a human-centric purpose. These measures can be categorised using existing frameworks, such as environmental, social and governance (ESG) criteria, AI design phases, or their intended audience.

To further illustrate how theoretical principles can be translated into practical applications, several concrete measures can be considered. These include reducing the high resource demands of AI model training and minimising the carbon and climate footprint of AI system design, as well as building a comprehensive understanding of the context in which the design operates and the power dynamics at play. Additionally, efforts should focus on mapping unpaid or exploited labour in the supply chain, testing training data for diversity measures, and ensuring explicit consent for

data collection and usage. Open-source alternatives should be elaborated to improve accessibility and transparency, while mechanisms must be established to challenge the reinforcement of harmful stereotypes. Finally, implementing regular maintenance measures and cycles will ensure continuous monitoring and control of undesired outcomes.

Ensuring AI systems meet societal demands and demonstrate implementation potential

Beyond designing purposeful AI systems, it is crucial to ensure their practical applicability. This involves convincing powerful stakeholders of the value of redistributing power among more diverse groups. For instance, governments must enact stronger protections against discrimination targeting marginalised groups, organisations need to prioritise hiring a more diverse workforce, and engineers should focus on inclusive design.

Implementing FAI in practice involves both mitigating potential risks and optimising outcomes. For example, FAI can lead to higher-quality outputs, such as increased accuracy rates for female users or marginalised communities. Business leaders increasingly recognise that responsible AI implementation is not only socially responsible, but also profitable in the long term, making it a compelling business case for investors. Furthermore, adopting a structured governance approach can reduce AI failures, minimise the loss of consumer trust caused by irresponsible practices, and ultimately enhance societal confidence in AI systems (Gillam and Jain, 2024).

Ensuring AI systems are continuously developed to reflect changing societal influences

A foundational feminist tool, the theory-practice loop, helps identify the real-world effects of implemented measures. While concepts such as equity, freedom and justice are inherently difficult to quantify, the nine principles outlined by D'Ignazio and Klein (2024) provide a framework to evaluate power dynamics and uncover biases within AI systems. For instance:

- checking if a system relies on **binaries and hierarchies**, such as binary gender categories, which may introduce gender bias and exclude individuals identifying as non-binary;

- assessing whether **pluralism** is embraced by valuing not just logical-mathematical intelligence, but other forms of intelligence such as interpersonal, bodily-kinaesthetic or naturalist intelligence (Onuoha and Nucera, 2023);
- ensuring that **consent** is explicitly obtained for the use of data in specific **contexts**, thereby preventing the exploitation of marginalised groups.

Other guiding principles include examining and challenging power structures, elevating emotion and embodiment, making labour visible and considering environmental impact (Klein and D'Ignazio, 2024).

This ongoing development aspect sets FAI apart from traditional responsible AI frameworks. While both share values like accountability and transparency and aim to create a positive societal impact, their underlying approaches differ. Responsible AI typically assumes that injustices stem from a few “bad actors” or groups, focusing on implementing ethical safeguards. In contrast, FAI critically examines systemic power structures as the root cause of inequities, aiming to address these structural dynamics at their core (D'Ignazio and Klein, 2020; Gillam and Jain, 2024).

Feminist AI in practice

FemAI: fostering equity at a legislative level

FemAI has petitioned the EU to incorporate feminist perspectives in the EU AI Act to protect MUUP from a regulatory standpoint (Wudel, Wedel and Gengler, 2023). However, the resulting EU AI Act reflects a balancing act that may compromise aspects of human rights protection (Wedel, Hazim and Wudel, 2024). The protection of environmental and societal sustainability is for instance considered as an optional code of conduct (European Parliament and Council of the European Union, 2024b).

Through a participatory approach, FemAI ensures the integration of feminist methodology by using online roundtables as expert focus groups. In this context, strategies have been developed to influence corporations via legislation. While the non-binding nature of the current framework presents challenges for the implementation of FAI, it also opens up opportunities for future improvement. The focus now shifts to assessing the impact of this code of conduct on promoting more equitable AI. Creative approaches are being explored to encourage policymakers to adopt a mandatory clause, moving closer to a more inclusive and impactful regulatory framework.

MIRA: reducing power inequalities in healthcare

FemAI has interviewed the CEO and team of a health tech start-up to share a concrete example of how AI tool providers can align with feminist values in practice. In

doing so, FemAI seeks to remind tech entrepreneurs of their power to design the future. As this paper aims to bridge the gap between principle and practice, the authors have decided to use a real-world example illustrating the technological advancement from our work in practice, in contrast to peer-reviewed sources which tend to be outdated after publication.

Responsible AI only works effectively when all key stakeholders are involved. A key to success for a feminist vision of AI is directly correlated with close cooperation amongst research, civil society and tech companies to directly influence AI developments before their deployment. MIRA Vision will be showcased as one potential example of how FemAI envisions tech tools that serve humanity.

The MIRA Vision project (MIRA) is a resource-efficient AI diagnostic platform committed to data privacy, diversity and social justice. By utilising synthetic, privacy-preserving image data, MIRA ensures secure and equitable access to diagnostic health data for all user groups. Its approach reduces hierarchies and barriers in healthcare, fostering societal sustainability and democratising medical knowledge.

MIRA adopts a multi-stakeholder, iterative development process, focusing on synthetic data training to address inequities while minimising resource demands. By avoiding deep learning algorithms, MIRA reduces costs, adapts quickly to evolving datasets, and remains accessible in resource-limited settings. This low-resource model also brings environmental advantages.

These innovations result in lower training costs and faster processing, enabling healthcare solutions to reach broader populations. To address inequities and make them measurable, MIRA analyses mortality rates, side effects and treatment outcomes across gender, age and ethnicity. This data-driven approach helps uncover systemic disparities, ensuring technical efficiency is aligned with social inclusivity. By bridging gaps in healthcare delivery, MIRA accelerates access and tackles critical inequalities.

Conclusion

FAI offers a critical lens for addressing the social, ethical and structural issues embedded in AI systems. By applying intersectional feminist principles, FAI seeks to ensure that AI systems are not only inclusive, but actively work to dismantle historical power imbalances. This paper underscores the importance of theory-practice feedback loops, continuous development and stakeholder engagement in building AI systems that foster equity and justice. The integration of FAI into both legislative frameworks and business practices will be essential in advancing socially responsible AI and creating a more equitable technological future that provides benefits for humanity.

References

- Arora, P. and Chowdhury, R. (2021) "Cross-Cultural Feminist Technologies", *Global Perspectives*, 2(1), p. 25207. Available at: <https://doi.org/10.1525/gp.2021.25207>.
- Atanasoski, N. and Vora, K. (2019) "Epilogue: On Technoliberal Desire, Or Why There Is No Such Thing as a Feminist AI", *Surrogate Humanity*. Durham, NC: Duke University Press, pp. 188–196. Available at: <https://doi.org/10.1515/9781478004455-009> (Accessed: 7 October 2024).
- D'Ignazio, C. and Klein, L.F. (2020) *Data Feminism*. Cambridge, MA: MIT Press. Available at: <https://doi.org/10.7551/mitpress/11805.001.0001>.
- European Parliament and Council of the European Union (2024a) Article 3: Definitions | EU Artificial Intelligence Act. Available at: <https://artificialintelligenceact.eu/article/3/> (Accessed: 12 November 2024).
- European Parliament and Council of the European Union (2024b) Article 95: Codes of Conduct for Voluntary Application of Specific Requirements | *EU Artificial Intelligence Act*. Available at: <https://artificialintelligenceact.eu/article/95/> (Accessed: 10 October 2024).
- Feminist Internet and Young, J. (2020) "Feminist Design Tool: Defensible Decision Making for Interaction Design and AI". Available at: https://ugc.futurelearn.com/uploads/files/16/b0/16b088ad-6145-45eb-b5d8-3753a41b4b88/2-10_Feminist-DesignTool_2.0.pdf (Accessed: 12 November 2024).
- Ferguson, K.E. (2017) "Feminist Theory Today", *Annual Review of Political Science*, 20(1), pp. 269–286. Available at: <https://doi.org/10.1146/annurev-polisci-052715-111648>.
- Gengler, E. J., Wedel, M. and Hazim, A. (2023) "Power Imbalances in Society and AI: On the Need to Expand the Feminist Approach".
- Gillam, C. and Jain, D. (2024) *Responsible AI Playbook for Investors 2024*. Geneva: World Economic Forum. Available at: <https://www.weforum.org/publications/responsible-ai-playbook-for-investors/> (Accessed: 13 October 2024).
- Hill Collins, P. (2010) "The New Politics of Community", *American Sociological Review*, 75(1), pp. 7–30. Available at: <https://doi.org/10.1177/0003122410363293>.
- Klein, L. and D'Ignazio, C. (2024) "Data Feminism for AI", *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency (FAccT '24)*. Rio de Janeiro, Brazil: ACM, pp. 100–112. Available at: <https://doi.org/10.1145/3630106.3658543>.
- Onuoha, M. and Nucera, D. (2023) *People's Guide to AI*. Edited by Alanna Maldonado.
- Purkayastha, B. (2012) "Intersectionality in a Transnational World", *Gender & Society*, 26(1), pp. 55–66. Available at: <https://doi.org/10.1177/0891243211426725>.
- Toupin, S. (2024) "Shaping Feminist Artificial Intelligence", *New Media & Society*, 26(1), pp. 580–595. Available at: <https://doi.org/10.1177/14614448221150776>.
- Wedel, M., Hazim, A. and Wudel, A. (2024) "Europäische KI-Regulierung: Auf der Suche nach verbindlichen Ansätzen für Nachhaltigkeit und Inklusion", *Data Sharing – Datenkapitalismus by Default? Nomos Verlagsgesellschaft mbH & Co. KG*, pp. 141–162. Available at: <https://doi.org/10.5771/9783748940173-141>.
- Wudel, A., Wedel, M. and Gengler, E. (2023) "A Feminist Vision for the EU AI Act". Available at: <https://alexandra-wudel.com/wp-content/uploads/2023/10/A-feminist-vision-for-the-EU-AI-Act.pdf> (Accessed: 6 November 2024).

About the Authors

FemAI is a responsible AI start-up certifying AI systems designed to benefit humanity. Established in 2022 in Berlin, Germany, we started as a **research-focused think tank**, advising over 25 organisations on their ethical AI policies. Building on these insights, we developed feminist frameworks to evaluate AI tools, companies, courses and workshops – focusing on the people most often overlooked in the AI age. At our core, FemAI uses **intersectional feminism** as a research method to illuminate the black box of AI, ensuring accuracy and equity in the technologies shaping our future. Co-creation and cross-sector collaboration are at our heart, allowing us to close the gaps between policymaking, businesses, research and civil society.

Alexandra Wudel is a tech entrepreneur and researcher focused on shaping an inclusive future for AI. As the CEO of FemAI, she was named “AI Person of the Year 2024”, recognised as one of 25 global experts on ethical considerations for AI, and awarded the title “35 under 35 Young Leaders”. Alexandra has contributed to the development of ethical AI guidelines for the German Federal Foreign Ministry, the UN and the German Parliament. Working with AI tool providers, she supports private organisations in integrating societal and environmental responsibility into AI projects and strategies. Together with her team, Alexandra develops FAI frameworks to bridge the gap between principles and practice.

Anna Ehrenberg is part of the FemAI team. She works as a research assistant at FemAI, actively shaping the organisation’s understanding of FAI. With her background in data science and process optimisation, she has co-authored the design of an FAI approach intended to be practically implemented in AI tools from the outset.

What is Feminist AI?

This paper explores Feminist Artificial Intelligence (FAI), a framework leveraging intersectional feminism to address biases and inequities in AI systems. FAI emphasises interdisciplinary collaboration, systemic power analysis and iterative theory-practice loops. By embedding feminist values – equity, freedom and justice – FAI seeks to transform AI development, ensuring inclusivity and social sustainability. Practical applications include initiatives such as FemAI's advocacy for feminist perspectives in the EU AI Act and the MIRA diagnostic platform, which aligns AI tools with social justice goals. FAI marks a critical departure from traditional AI by tackling structural inequalities and promoting accountable and equitable AI solutions.

Further information on the topic can be found here:

➤ fes.de