

Anastasia Karagianni
January 2025

The EU Artificial Intelligence Act through a Gender Lens

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 978-3-98628-690-3

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

Anastasia Karagianni
January 2025

The EU Artificial Intelligence Act through a Gender Lens

Content

Summary	3
Introduction	3
Background: AI systems that have been used and their implications on gender equality and non-discrimination	4
A gender-responsive text analysis of the AI Act	5
Article-by-article analysis of the AI Act from a feminist perspective	6
Article 5 on vulnerability	6
Article 6 on harmonisation	6
Article 40(1) on standardisation	7
Article 43 on conformity assessment	8
Article 27 on the Fundamental Rights Impact Assessment (FRIA) and Gender Impact Assessment	9
Conclusion	10
Recommendations	11
References	12

Summary

This paper provides a feminist analysis of the European Union's Artificial Intelligence Act (hereinafter the "AI Act"), assessing its capacity to address gender inequalities and structural power imbalances in AI systems. Drawing on feminist theories, the paper evaluates the AI Act's limitations in mitigating gender biases that might disproportionately impact marginalised groups, particularly women of colour and women from marginalised communities. Through case studies in recruitment and employment, healthcare, border management control and predictive policing in domestic violence cases, the report highlights how AI applications can reinforce gender disparities. A detailed examination of specific provisions within the AI Act reveals critical gaps in addressing systemic discrimination and bias in AI governance. To promote a more equitable AI landscape, the report recommends integrating intersectional, feminist-informed revisions that prioritise interdisciplinarity, collective instead of individual approaches and strong oversight mechanisms that are oriented by human rights – including feminist – values. The proposed recommendations focus on strengthening the AI Act's framework to better safeguard marginalised communities and to ensure a regulatory approach that reflects the diverse experiences of all individuals.

Introduction

This paper examines the need for feminist-informed AI frameworks to address diverse socio-technical impacts and counter the risk of AI reflecting biases favouring white, cisgender, able-bodied men. Drawing on objectification theory (Frederickson and Roberts, 1997)¹, the report investigates how AI systems may replicate societal norms that reduce women to physical appearances, reinforcing male-dominated perspectives, as noted by MacKinnon (MacKinnon, 2013), and perpetuating biases seen in media portrayals, thereby amplifying sexism and gender-based inequalities in digital technologies.

This paper argues that robust regulation of AI is critical, examining the AI Act through a feminist lens to identify how it interacts with EU laws, including the General Data Protection Regulation (GDPR), equality and non-discrimination laws and consumer protection legislation. Recommendations focus on interpreting the AI Act to strengthen provisions on inclusivity, diversity, transparency and accountability, ensuring AI systems are regulated to prevent gender biases and protect marginalised groups.

¹ Objectification theory is a psychological framework that explains how living in a culture that sexually objectifies the female body affects women's mental and physical health. The theory was first developed by Barbara L. Fredrickson and Tomi-Ann Roberts in their 1997 paper, "Objectification Theory: Toward Understanding Women's Lived Experiences and Mental Health Risks". See Fredrickson, B. L., & Roberts, T. A. (1997). Objectification Theory: Toward Understanding Women's Lived Experiences and Mental Health Risks. *Psychology of Women Quarterly*, 21(2), 173-206.

Background: AI systems that have been used and their implications on gender equality and non-discrimination

AI technologies, if not designed with comprehensive oversight, can inadvertently perpetuate existing societal biases, leading to discriminatory impacts against women and marginalised communities. The following case studies underscore this urgent need for regulatory measures that ensure AI development and deployment practices prioritise inclusivity, gender sensitivity and transparency to mitigate the potential for systemic bias and discrimination.

There have been a number of AI systems already being used across different domains. In employment, Amazon's AI recruitment tool (Dastin, 2018) – intended to streamline hiring – unintentionally favoured male candidates, as its algorithms were trained on predominantly male résumés, leading to gender-biased selection criteria and eventually necessitating its discontinuation.

A more recent example of a discriminatory AI system is the case of Deliveroo in Italy (Zuiderveen Borgesius, 2018). The Court ruled against Deliveroo's rider-ranking algorithm, citing discrimination. The algorithm was used to determine the "reliability" of a rider. If the rider failed to cancel a pre-scheduled shift at least 24 hours in advance, the algorithm would begin to downgrade the rider, impacting their ability to gain future shifts. The Court found the algorithm violated local labour laws because it did not distinguish between legally protected reasons for withholding labour, such as sickness, emergency or exercising their protected right to strike, versus unprotected reasons for failing to be available. This means that in the case of a woman experiencing severe menstrual pain, she may be unable to cancel her shift 24 hours in advance, as it is not always possible to predict the onset or severity of her symptoms.

In healthcare, AI applications in obstetrics for the prediction of preterm birth and other complications (Belciug and Iliescu, 2023) or in gynaecological cancer detection – which is conducted based on electronic health records that often rely on binary gender frameworks, failing to accurately accommodate transgender and non-binary patients (Taylor and Bryson, 2016) – can result in misclassification, inaccuracies and inadequate care.

In border management control, the iBorderCtrl pilot project, an EU initiative for AI-driven border security,

used facial recognition and lie detection technologies to assess traveller credibility. However, this system has raised concerns about potential gender and racial biases in AI, which could lead to discriminatory treatment of travellers (Rosario, 2023). Specifically, the algorithms could misinterpret women's facial expressions, reflecting societal biases that perceive women as more emotional or less credible, thereby increasing the likelihood of false-positive detections among female travellers. The project underscored the need for rigorous bias mitigation, transparency and accountability in AI deployment within border security, as failure to address these issues could compromise fairness, inclusivity and human rights protections at the border.

The VioGen case – which involves the implementation of a system for assessing and managing gender-based violence risks used by the Spanish Ministry of the Interior – highlights significant concerns regarding gender bias in AI-driven decision-making tools in predictive policing (Ruibal Pérez, 2019). Despite its aim to improve responses to domestic violence and take protection measures, the system has been criticised for relying on algorithms that inadvertently reinforced existing gender stereotypes and biases, particularly in the assessment of victims and offenders¹. The risk assessment criteria do not adequately account for the complexities of individual cases, leading to misclassification and potentially harmful outcomes for victims/survivors, particularly women. To the extent of gender-based violence, the rapid advancement of generative AI tools, especially in creating deepfakes, has raised legal concerns, as these tools enable the production of realistic, non-consensual and sexualised depictions of individuals, constituting a modern form of AI-generated gender-based violence (Karagianni and Doh, 2024).

All these cases underscore the urgent need for policymakers to ensure that AI applications both in the public and private sector – for example, in recruitment and employment, healthcare, border management control and predictive policing – are designed with a strong emphasis on gender sensitivity, transparency and accountability, to mitigate the risk of perpetuating bias and to promote equitable protection for all individuals affected by gender-based violence.

¹ See, <https://eticasfoundation.org/?audit-spotlight=the-adversarial-audit-of-viogen>.

A gender-responsive text analysis of the AI Act

On 22 January 2024, the pre-final text of the AI Act was released, receiving unanimous approval from the EU Council on 2 February 2024. This legislation follows the foundational framework established by the White Paper on Artificial Intelligence (European Commission, 2020), which aimed to protect fundamental rights and prohibit AI practices that conflict with EU values and fundamental rights under EU law. A critical examination of the AI Act reveals significant gaps in addressing gender equality, which are particularly relevant in the context of algorithmic fairness and data protection.

Notably, the earlier drafts of the AI Act did not reference gender equality, largely due to opposition from a Member State, Poland, regarding the inclusion of the term (Stolton, 2020). This lack of reference highlights a broader issue within EU regulations, as gender has never been a focus in AI policy discussions, which were centred on AI definitions.

In recognising the risks associated with algorithmic bias, the AI Act includes multiple references to “non-discrimination” in Recitals 27, 28, 31, 48, 58, 60 and 142 AI Act, as well as in Article 77(1) AI Act, which empowers authorities to protect fundamental rights. The term “discrimination” is mentioned in Recitals 56, 57, 58, 67, 70, 80 and 110 AI Act, and in Articles 10(2) (f) AI Act and Annex IV (3) AI Act. While the AI Act acknowledges the importance of non-discrimination, references to “gender equality” are limited to Recitals 27 and 48 AI Act, along with one mention in Article 95(2)(e) AI Act, which requires the AI Office and member states to assess and mitigate the negative impacts of AI systems on vulnerable groups, including those pertaining to gender equality. Although the term “women” appears in Recitals 56 and 57 AI Act to acknowledge discrimination patterns against women and individuals based on sexual orientation, the text notably omits more inclusive terminology such as “non-binary”, “transgender”, “intersex” and “gender non-conforming people”.

This linguistic analysis of the AI Act reveals its predominantly gender-neutral language, which avoids gender-specific terms. However, to substantiate a genuine commitment to gender equality, the Act would benefit from adopting more explicit gender-responsive language. Gender-neutral language may in-

advertently overlook the unique challenges faced by marginalised groups, while gender-responsive language actively promotes equality and addresses the specific needs of diverse gender identities.

In conclusion, the AI Act represents a significant step towards regulating artificial intelligence within the EU framework, yet it falls short in adequately addressing gender equality and the risks of bias inherent in AI systems. As policymakers move forward, it is imperative to enhance the Act with more comprehensive and inclusive language that explicitly addresses the intersection of gender and AI, ensuring that regulations not only prevent discrimination, but also actively promote gender equity in all dimensions of AI development and deployment. The integration of gender-responsive provisions will be crucial in safeguarding fundamental rights and advancing the EU’s commitment to equality in the rapidly evolving landscape of artificial intelligence.

Article-by-article analysis of the AI Act from a feminist perspective

Article 5 on vulnerability

Article 5 of the AI Act delineates AI practices deemed unacceptable owing to their potential to jeopardise safety and fundamental rights, explicitly prohibiting systems such as social scoring (Article 5(1)(c) AI Act) and manipulative techniques targeting *vulnerable* individuals (Article 5(1)(a) and (b) AI Act). *Vulnerability* is a concept rooted in both computer science and social sciences (Praveen Kumar, 2020). In computer science, vulnerability is defined as a security flaw, glitch or weakness found in software code that could be exploited by an attacker². In social sciences, it highlights the reliance on external support that arises from diminished self-determination, as seen for instance in the harmful portrayal of Indigenous women as sexualised and inherently vulnerable, echoing colonial and patriarchal narratives (Agloinga, 2021)³.

While Article 5 prohibits the use of biometric categorisation systems involving sensitive characteristics such as race and sexual orientation in public spaces, it notably omits gender from this list, raising concerns about the protection of gender in contexts of social behaviour and performativity. Despite recommendations from the European Data Protection Board to include gender as a sensitive category (EDPS and EDPB, 2021), this has not been integrated into the AI Act. This oversight underscores a critical gap in ensuring gender equality in AI systems, necessitating an intersectional (Crenshaw, 2019)⁴ approach that acknowledges and addresses the complex layers of discrimination and oppression inherent in these technologies.

Article 6 on harmonisation

Article 6 of the AI Act delineates the requirements for high-risk AI systems, mandating compliance with

standards for safety, transparency and accountability to promote a unified approach to AI governance across EU member states. This requirement for harmonised standards refers to the technical standards that provide clear methods to meet legal requirements in the preparation stage of the AI system. Following European harmonised standards, as listed in the EU's Official Journal, gives a legal assumption of compliance with regulations⁵. This helps create fair conditions for designing and developing AI systems, especially benefiting small and medium-sized AI businesses. Recital 121 AI Act underscores the necessity of harmonised standards across the European Union, which promotes a consistent application of the AI Act's provisions among member states.

By harmonising national laws with EU legislation, such as the EU Equality and Non-Discrimination Law, GDPR, Digital Services Act (DSA), Digital Markets Act (DMA), and General Product Safety Regulation (Regulation (EU) 2023/988), the AI Act seeks to create a cohesive legal framework that safeguards fundamental rights while ensuring consistency and fairness in AI practices. From a feminist perspective, it is critical to analyse how these legislative efforts address issues of gender equality, inclusivity and social justice, particularly for marginalised communities.

Feminist scholars (Theilen, 2021) contend that existing laws such as the GDPR must incorporate stronger protections for individuals, especially women, facing vulnerabilities related to privacy breaches and surveillance, as seen in domestic violence situations. This includes enhancing controls over personal data, particularly regarding location tracking and consent (Sovacool et al., 2021), to empower individuals against data misuse, such as image-based sexual abuse and non-consensual deepfakes⁶.

² See the First Draft General-Purpose AI Code of Practice, published by the European Commission, p. 17, where it states that vulnerability discovery recognised as a cyber offence is listed in the taxonomy of systemic risks (<https://digital-strategy.ec.europa.eu/en/library/first-draft-general-purpose-ai-code-practicepublishedwrittenindependent-experts>).

³ This intersection of nudity and vulnerability is deeply rooted in colonialism, exoticism and patriarchy, and continues to affect how indigenous women are viewed and treated today, while it raises critical questions on how AI technologies are designed and deployed. See Agloinga, C. M. (2021). *Indigenous Women: Violence, Vulnerability and Cultural Protective Factors* (Doctoral dissertation, Walden University).

⁴ Intersectional feminism takes into account the many different ways each woman experiences discrimination. Crenshaw first used the term "intersectionality" to refer to the double discrimination of racism and sexism faced by Black women, critiquing the "single-axis framework that is dominant in antidiscrimination law". See Crenshaw, K. (2019). 'Difference' through intersectionality. *Dalit Feminist Theory* (pp. 139–149). Routledge India.

⁵ See also, European Joint Research Centre of the European Commission (JRC), 2024. Harmonised Standards for the European AI Act. <https://publications.jrc.ec.europa.eu/repository/handle/JRC139430>.

⁶ Abusive partners might use digital surveillance tools, such as Global Positioning System (GPS) tracking or spyware on phones and computers, to control and monitor the movements and communications of pregnant women. See Sovacool, B., Furszyfer-Del Rio, D. D., & Martiskainen, M. (2021). Can prosuming become perilous? Exploring systems of control and domestic abuse in the smart homes of the future. *Frontiers in Energy Research*, 9, 765817.

To this extent, the DSA's provisions against illegal content must similarly address tech-facilitated gender-based violence advocating for accountability in content moderation that disproportionately affects marginalised users, including women of colour (Karagianni and Doh, 2024).

Furthermore, the concept of intersectionality, as articulated by Kimberlé Crenshaw (Crenshaw, 2019), underscores the need for EU harmonisation efforts to consider how various forms of discrimination intersect in AI. The intersectionality framework underscores that individuals do not experience bias in isolation; rather, their experiences are shaped by the confluence of their gender, race, socioeconomic status and other identity factors. In the context of bias auditing, an intersectional approach allows for a deeper understanding of how AI systems may disproportionately impact women from various backgrounds, including women of colour, LGBTQIA+ individuals, and those with disabilities. For example, an AI hiring tool may not only exhibit gender bias against women, but could also reflect racial biases that further disadvantage women of colour.

The European Union's Directive on combating violence against women and domestic violence acknowledges the importance of intersectionality in Recitals 6 and 71. The Directive includes provisions that recognise the diverse and intersecting identities of individuals, aiming to ensure that measures against gender-based violence are inclusive and effective across various social groups. This approach is designed to address the unique challenges faced by individuals who may experience multiple forms of discrimination simultaneously.

To conclude, a feminist approach to harmonisation must avoid homogenisation, ensuring that legal frameworks are adaptable to the diverse social, economic and cultural contexts of EU member states while prioritising the protection of marginalised communities. Embedding feminist principles, such as intersectionality, into the AI Act is essential for advancing gender equality and inclusivity in AI governance.

Article 40(1) on standardisation

The AI standardisation process is essential for establishing comprehensive guidelines, technical specifications and best practices to ensure that AI systems are safe, reliable, transparent and ethically designed. As articulated in Article 40 and Recital 121 of the AI Act, standardisation plays a pivotal role in fostering innovation while minimising risks associated with bias and discrimination.

AI standardisation involves the development and implementation of norms and technical standards that ensure

AI systems operate ethically and effectively. Various entities, including European Standards Organisations (ESOs), National Standards Bodies (NSBs) and the European Commission, are working together to establish frameworks regulating AI. Key areas of focus include human-centred AI, security and privacy, and data governance. This analysis will specifically highlight the importance of explainability and accountability within AI systems, which are critical to upholding fundamental rights.

In May 2023, the European Commission issued a formal standardisation request, which CEN-CENELEC accepted. Work on these standards is ongoing, driven by input from diverse stakeholders. The request emphasised the need to include representatives from various sectors and organisations. However, achieving consensus on important issues has been difficult, leading to a slower-than-expected process, even by the standards of those involved.

It is crucial that the AI Act standards prioritise addressing the risks AI poses to people's health, safety and fundamental rights. This is a new focus for AI standards, as most international standards have mainly sought to protect the goals of organisations using AI. Another important requirement is that oversight of AI systems must produce verifiable results. This means testing must show that measures to reduce risks are effective, with human involvement when necessary.

Explainability, or interpretability, are key demands of civil society organisations that strengthen the capacity of individuals to understand the outcomes and decisions generated by AI systems. From a feminist perspective, explainability extends to fostering "response-ability", enabling individuals – particularly those from marginalised communities – to critically assess and engage with AI systems. Feminists advocate for enhanced explainability in domains such as social media content moderation, which often operates in opaque ways that reinforce normative gender roles, disproportionately targeting women and their bodies. For example, platforms such as Tumblr (Duguay, 2018) have banned images depicting female-presenting nipples, illustrating how content moderation practices can perpetuate systemic inequalities. Initiatives such as Moderation Arcana⁷ aim to improve content moderation algorithms by incorporating user-centred research, thus promoting equality.

Alternatively, augmented-reality (AR) beauty filters used on social media deploy subliminal techniques beyond users' consciousness and manipulative or deceptive techniques, with the objective or effect of materially distorting the users' self-appreciation and beauty stand-

⁷ See, <https://www.themoderationarcana.com/>. In 2022, an algorithmic artist and transgender individual, Ada Ada Ada, attempted to test how algorithms perceive gender. By using her own transgender body, Ada Ada Ada found several methods for tricking gender recognition technology into seeing a gender different from the initial judgement on social media platforms. <https://ada-ada-ada.art/linktree>.

ards, thereby causing them significant psychological harm (Doh et al., 2024). The same transparency rules are applied for deployers and distributors and for the same reasons, they should prohibit the distribution of AR filters in their platforms. These techniques have a harmful effect on users' mental autonomy, and therefore transparency obligations and cultural diversity should be safeguarded. In these terms, explainability standards should be strengthened, especially when social media influencers promote a product – such as a moisturiser – using a beauty filter, where consumer law applies⁸.

Inclusive standardisation mandates the active involvement of diverse stakeholders, including women of colour and other marginalised groups, to ensure that the perspectives of those most affected by AI technologies are incorporated. Gender-aware standards should include measures to address gender bias and inequality, such as mandatory bias audits and the utilisation of diverse datasets. However, there is a significant risk of tokenism, where the involvement of diverse stakeholders is superficial and fails to result in substantive changes promoting gender equity. This challenge is exacerbated by the AI standardisation process being predominantly shaped by powerful corporations and governments, which may obstruct genuine feminist efforts to cultivate inclusivity and equity in AI development and regulation. To overcome these barriers, it is essential to prioritise collaborative approaches that actively engage marginalised voices, ensuring that the evolving landscape of AI remains equitable and just for all stakeholders involved.

Article 43 on conformity assessment

High-risk AI systems are mandated to undergo conformity assessments – during the evaluation stage of the AI systems (Baral, 2024) – to ensure adherence to established safety and ethical standards, as stipulated in Article 43 AI Act. This process verifies that the system complies with the necessary requirements before it can be marketed. Articles 6 and 43 AI Act create a comprehensive regulatory framework for high-risk AI systems, establishing a governance structure aimed at overseeing compliance and enforcement. Article 6 emphasises the importance of safety, transparency and accountability in AI technologies, while Article 43 facilitates cooperation and guidance among EU member states. This dual approach promotes responsible AI development and use while safeguarding the rights and interests of individuals.

AI conformity assessment involves evaluating whether an AI system meets specific standards, regulations and ethical guidelines prior to deployment. These assessments typically focus on ensuring that AI systems are safe, reliable, fair and devoid of bias. However, from a feminist perspective, it is crucial to address gender bias, inclusivity and power structures, particularly given the systemic inequalities that may be perpetuated if assessments lack an intersectional lens.

A prominent feminist concern regarding AI is the reinforcement of gender bias through biased algorithms and data. AI systems often rely on historical data that reflect existing social inequalities (Hendrickx, 2024). Consequently, conformity assessments should require comprehensive bias audits that extend beyond identifying overt discrimination. Such audits must also pinpoint subtle, structural biases against women, particularly women of colour, LGBTQIA+ individuals, and other marginalised groups. For instance, AI applications in recruitment or healthcare may disadvantage women by failing to adequately represent their experiences and needs in training datasets.

Intersectional data analysis in conformity assessments is therefore essential. Gender bias should not be evaluated in isolation. Assessments must consider how gender intersects with race, class, disability and other identity factors to ensure that AI systems do not disproportionately harm marginalised communities. To achieve this, conformity assessments should involve diverse teams that reflect a range of genders, races and social backgrounds. This diversity ensures that AI systems are scrutinised through multiple lenses, minimising bias and enhancing fairness. However, with women holding only 22% of AI stakeholder positions, considerable progress is still required to achieve genuine diversity in the field⁹. For example, Recital 165 advocates the participation of varied stakeholders in shaping AI regulations to enhance societal benefits and address potential risks effectively. When it comes to the diversity of AI terms, Recitals 4, 27, 80 and 165 AI Act underscore the importance of diversity in the development and deployment of AI systems. Even so, there are still fewer women than men platform workers, for instance, although this number increased after Covid-19 as platform work facilitated a work-life balance (EIGE, 2021).

Furthermore, it is critical to establish independent oversight bodies to ensure that AI conformity assessments are free from conflicts of interest and genuinely committed to

⁸ In February 2021, the Advertising Standards Authority (ASA) ruled that filters used in adverts should not mislead consumers about the effect of products. See <https://www.asa.co.uk/news/2021/02/24/beauty-filters-face-legislation-to-protect-mental-health/a-65681068>.

⁹ See https://www.turing.ac.uk/sites/default/files/2021-03/where-are-the-women_public-policy_full-report.pdf. Recital 150 also highlights the need for broad stakeholder involvement, particularly from marginalised communities, in developing AI regulations, while Recitals 165 and 173, alongside Articles 57 and 58 AI Act, underscore the critical importance of participatory governance in the regulation of AI systems. Furthermore, Recital 165 emphasises the need for active engagement from various stakeholders, including civil society, in the development and implementation of AI policies, ensuring that diverse perspectives are considered in decision-making processes. Recital 173 further supports this by highlighting the role of public consultation in fostering transparency and accountability within AI governance. Articles 57 and 58 establish mechanisms for stakeholder involvement in conformity assessments and oversight, reinforcing the commitment to inclusive participation that upholds democratic values and addresses the interests of all affected communities.

addressing issues of bias and harm. These bodies should have the authority to enforce compliance and hold AI developers accountable for any violations, thereby reinforcing the integrity of the AI assessment process and promoting equitable outcomes for all stakeholders.

Article 27 on the Fundamental Rights Impact Assessment (FRIA) and Gender Impact Assessment

Among the solutions proposed in the AI Act to protect fundamental rights endangered by high-risk AI systems are the risk management system (RMS) outlined in Article 9 and the Fundamental Rights Impact Assessment (FRIA) detailed in Article 27(1)¹⁰. However, questions arise regarding their adequacy in preventing gender-based discrimination and promoting gender equality within the context of AI. While the RMS and FRIA represent novel measures, risk management and impact assessments are not new constructs in technology regulation, but have historically emerged to address uncertainties associated with technological advancements.

Risk management is defined as the process of addressing potential negative events, while impact assessments evaluate both positive and negative consequences of initiatives on societal concerns, such as fundamental rights (Kloza et al., 2021). However, the existing literature lacks a comprehensive analysis of their differences and practical application, particularly in AI contexts (Kloza, 2018). For instance, defining what constitutes a risk to rights such as non-discrimination (Uuk et al., 2024) involves multiple conceptualisations, complicating risk measurement – a subjective process often influenced by gendered assumptions (Demetzou, 2019; Van Dijk et al., 2016). Moreover, the relationship between the RMS, FRIA and Gender Impact Assessments (GIAs), which specifically address impacts on gender equality, remains unclear.

The tools specified in Articles 9 and 27(1) of the AI Act may differ in scope and execution, yet both exhibit limitations in achieving gender equality. This is rooted in the normative framework of the AI Act, which often reflects a narrow understanding of gender and prioritises product safety over broader fundamental rights protections. To address these deficiencies, incorporating GIAs into the risk management systems and fundamental rights impact assessment is essential, reinforcing the principle that “women’s rights are human rights” (Bunch, 1990). Without this integration, the risk management systems and fundamental rights impact assessment risk protecting only the rights of a standard legal liberal persona, typically male and often white and able-bodied.

As feminist scholars indicate, fundamental rights and their assessment methods can function as either emancipatory or oppressive tools (Francs and Smith, 2021). GIAs highlight how development initiatives affect individuals differently based on their gender. They aim to uncover disparities caused by entrenched structural inequalities (Götzmann and Bainton, 2021).

Examples of GIA:

- Identification of potential impacts on individuals based on their gender, such as women, men, non-binary, intersex, gender non-conforming people etc.
- Assessment of whether the AI system reinforces existing gender inequalities or helps reduce them.
- Examination of how gender intersects with other factors such as race, class, ethnicity, disability and age (intersectionality).
- Consideration of gender-specific rights under the European Convention on Human Rights, the UN Convention on the Elimination of All Forms of Discrimination Against Women (CEDAW) and other human rights treaties, such as reproductive rights in the case of AI deployment in obstetrics.
- Collection and analysis of gender-disaggregated data to understand different gender needs and experiences.
- Use of gender-sensitive indicators to measure impacts in areas such as labour market participation, health outcomes, social protection or access to justice.

However, conflating GIAs with risk management systems and fundamental rights impact assessment may lead to ineffective echoed activities, undermining the genuine protection of fundamental rights. Therefore, alternative solutions should be explored, such as the integration of multiple focused impact assessments (Kloza et al., 2021) or innovative methods within existing assessment frameworks (Calvi, 2022).

¹⁰ Until 2 November 2024, member states had to identify public bodies which supervise or enforce fundamental rights in relation to the use of high-risk AI systems referred to in Annex III and make the list publicly available. See European Commission, General-Purpose AI Code of Practice: List of Participating Organisations, published on 6 November 2024. <https://digital-strategy.ec.europa.eu/en/policies/ai-code-practice>.

Conclusion

The AI Act aims to protect fundamental rights but lacks strong measures for gender equality. While it references non-discrimination, explicit mentions of “gender equality” and diverse gender identities are minimal. The predominantly gender-neutral language may miss the challenges/voices of marginalised communities. To ensure equality for all, policymakers are urged to integrate gender-responsive provisions, addressing any prejudice in the language of the text and promoting equality and diversity across all dimensions of AI regulation and deployment.

Through a feminist analysis of the AI Act, it became apparent that the intersectional approach of gender-based discrimination in AI is more than necessary (Leavy, 2018). AI systems depend on data, causing data bias resulting from unrepresentative, incomplete or historically biased datasets and perpetuating existing inequalities. For this reason, any kind of user vulnerability emerging from systemic inequalities should be considered under diversity, inclusion and participation provisions. The AI Act –when it enters into force – should be harmonised with EU law. Once enshrined in EU law, it should tackle any form of gender-based discrimination. The standardisation process, which occurs during the preparation stage of an AI system, and conformity assessments, which take place during the evaluation stage of the AI system before its market release, can also provide adequate protection for the prevention and mitigation of gender bias. Yet deployers of AI systems should conduct a fundamental rights impact assessment even after the system is released. According to this feminist analysis, a GIA under the fundamental rights impact assessment is essential for comprehensive and efficient protection of gender equality in AI.

While the AI Act holds potential, it is not a panacea; rather, it must be developed with a critical awareness of existing barriers to women and LGBTQIA+ people and an acknowledgment, for instance, of gender-based violence and oppression. Interdisciplinary collaboration, rooted in feminist critical analysis of law, is essential to ensure that AI technologies foster equality and resist perpetuating existing inequalities, ultimately facilitating an inclusive and equitable digital landscape.

Recommendations

EU institutions are increasingly recognising the significance of intersectionality in policymaking. However, practical comprehension among those involved remains limited, resulting in a notable enforcement gap.

- A clear definition of intersectional discrimination and advocates for its integration into policymaking processes are needed to effectively address systemic inequalities.
- The principle of equal treatment must extend beyond mere prohibitions against discrimination based on sex or gender to encompass discrimination rooted in gender identity.
- Existing AI policies must not inadvertently re-marginalise specific groups, particularly women of colour.
- AI systems should adhere to principles of transparency, explainability, fairness and accountability, implementing measures such as mandatory audits to mitigate potential discrimination.
- Gender equality impact assessments as outlined in Articles 9 and 27(a) of the AI Act should be rigorously enforced.
- Inclusive participation of non-binary, intersex, transgender and gender non-conforming individuals in design and development teams is essential, as mandated in Article 69(2)(d).

Inclusive language:

- The interpretation of gender equality within the Act should be broadened to signal a commitment to an inclusive understanding of equality and non-discrimination. As articulated, the spectrum of equality must be viewed through a multifaceted lens that addresses intersectional discrimination.
- The explicit inclusion of transgender people within EU non-discrimination law, such as the Directive on combating violence against women and domestic violence, is required. Although the Court of Justice of the European Union (CJEU) has interpreted sex discrimination to encompass transgender people post-gender reassignment, there has been no similar recognition for

non-binary or intersex persons. This raises questions about the adequacy of EU non-discrimination law for those unable or unwilling to access gender-affirming healthcare.

- The definition of gender must acknowledge its complex and fluid dimensions, as recognised in reports from the Council of Europe and the Yogyakarta Principles. Such acknowledgment is vital to dismantling the patriarchal, heteronormative structures that oppress women and marginalise non-conforming identities.

References

- Agloinga, C. M. (2021). *Indigenous Women: Violence, Vulnerability and Cultural Protective Factors* (Doctoral dissertation, Walden University).
- Baral, A. K. (2024). Standards Development: The Starting Point of Conformity Assessment. In *Handbook of Quality System, Accreditation and Conformity Assessment* (pp. 1–31). Singapore: Springer Nature Singapore.
- Belciug, S., & Iliescu, D. Pregnancy with Artificial Intelligence. See also, Kumar, D. S., & Maniwaran, P. (Eds.). (2023). *Predicting Pregnancy Complications Through Artificial Intelligence and Machine Learning*. IGI Global.
- Bunch, C. (1990). Women's Rights as Human Rights: Toward a Re-vision of Human Rights. *Human Rights Quarterly*, 12, 486.
- Calvi, A. (2022). Gender, data protection & the smart city: Exploring the role of DPIA in achieving equality goals. *European Journal of Spatial Development* (EJSD) 19, 3 (2022), 24–47. DOI: <https://doi.org/doi:10.5281/zenodo.6539249>.
- Crenshaw, K. (2019). 'Difference' through intersectionality. *Dalit Feminist Theory* (pp. 139–149). Routledge India.
- Dastin, J. (2018). Insight – Amazon scraps secret AI recruiting tool that showed bias against women. *Reuters*.
- Demetrou, K. (2019). GDPR and the Concept of Risk: The Role of Risk, the Scope of Risk and the Technology Involved. *Privacy and Identity*, Eleni Kosta (ed.). Springer International Publishing, 137–154. DOI: https://doi.org/10.1007/978-3-030-16744-8_10. See also, Niels Van Dijk, Raphaël Gellert and Kjetil Rommetveit. 2016. A risk to a right? Beyond data protection risk assessments. *Computer Law and Security Review* 32, 2 (2016), 286–306. DOI: <https://doi.org/10.1016/j.clsr.2015.12.017>.
- Doh, M., Canali, C., & Karagianni, A. (2024). Pixels of Perfection and Self-Perception: Deconstructing AR Beauty Filters and Their Challenge to Unbiased Body Image. In *Proceedings of the 2024 ACM International Conference on Interactive Media Experiences* (pp. 349–353).
- Duguay, S. (2018). *Why Tumblr's ban on adult content is bad for LGBTQ youth*. The Conversation.
- EIGE (2021). Artificial intelligence, platform work and gender equality. EIGE.
- European Commission (2020). *White Paper on Artificial Intelligence: A European approach to excellence and trust*. European Commission.
- European Data Protection Supervisor (EDPS) & European Data Protection Board (EDPB) (2021). Joint opinion 5/2021 on the European Commission's proposal for a regulation on artificial intelligence (AI).
- European Joint Research Centre of the European Commission (JRC), 2024. Harmonised Standards for the European AI Act. <https://publications.jrc.ec.europa.eu/repository/handle/JRC139430>.
- Frans, L. and Smith, P. (2021). Feminist Philosophy of Law. *The Stanford Encyclopedia of Philosophy*.
- Fredrickson, B. L., & Roberts, T. A. (1997). Objectification Theory: Toward Understanding Women's Lived Experiences and Mental Health Risks. *Psychology of Women Quarterly*, 21(2), 173–206.
- Götzmann, N., & Bainton, N. (2021). Embedding gender-responsive approaches in impact assessment and management. *Impact Assessment and Project Appraisal*, 39(3), 171–182.
- Hendrickx, V. (2024). Women's rights in the age of automation. *KU Leuven blogpost*. <https://www.law.kuleuven.be/citip/blog/womens-rights-in-the-age-of-automation/>.
- Karagianni, A., & Doh, M. (2024). A feminist legal analysis of non-consensual sexualized deepfakes: contextualizing its impact as AI-generated image-based violence under EU law. *Porn Studies*, 1–18. <https://doi.org/10.1080/23268743.2024.2408277>.
- Kloza, D. (2018). Sketching the relationship between risk (management) and impact (assessment).
- Kloza, D., Van Dijk, N., Casiraghi, S., Vazquez Maymir, S. and Tanas, A. (2021). The concept of impact assessment. *Border Control and New Technologies*, Peter J. Burgess and Dariusz Kloza (eds.). ASP, Brussels, 31–48. DOI: <https://doi.org/10.46944/9789461171375.2>.
- Leavy, S. (2018). Gender bias in artificial intelligence: The need for diversity and gender theory in machine learning. *Proceedings of the 1st international workshop on gender equality in software engineering* (pp. 14–16).
- Praveen Kumar, E. (2020). A Review on Vulnerabilities to Modern Processors and its Mitigation for Various Variants. *Procedia Computer Science*, Vol. 215, pp. 91-97.
- Rosario, F. (2023). Gendered Borders: Unveiling the Impact of EU's Technological Fortress on Women and Queer Refugees. *Institut du Genre en Geopolitique*.
- Ruibal Pérez, J. Á. (2019). Protection of victims of violence against women in Spain. *Valahia University Law Study*.
- Sovacool, B., Furszyfer-Del Rio, D. D., & Martiskainen, M. (2021). Can prosuming become perilous? Exploring systems of control and domestic abuse in the smart homes of the future. *Frontiers in Energy Research*, 9, 765–817.
- Stolton, S. (2020). *Poland rejects Presidency conclusions on Artificial Intelligence, rights*. Euractiv.
- Taylor, E. T., & Bryson, M. K. (2016). Cancer's margins: trans* and gender nonconforming people's access to knowledge, experiences of cancer health, and decision-making. *LGBT Health*, 3(1), 79–89.
- Theilen, J. T. & Baur, A. & Bieker, F. & Ammicht Quinn, R. & Hansen, M. & González Fuster, G. (2021). Feminist data protection: an introduction. *Internet Policy Review*, 10(4). <https://doi.org/10.14763/2021.4.1609>.
- Uuk, R. & Gutierrez, C. I., & Lauwaert, L. et al. (2024). A Taxonomy of Systemic Risks from General-Purpose AI. Available at SSRN: <https://ssrn.com/abstract=>
- Van Dijk, N., Gellert, R., and Rommetveit, K. (2016). A risk to a right? Beyond data protection risk assessments. *Computer Law and Security Review* 32, 2 (2016), 286–306. DOI: <https://doi.org/10.1016/j.clsr.2015.12.017>.
- Zuiderveen Borgesius, F. (2018). Discrimination, Artificial Intelligence and Algorithmic Decision-Making. *Council of Europe*.

About the Author

Anastasia Karagianni is a [FARI Scholar](#) and Doctoral Researcher at the LSTS research group of Law and Criminology Faculty at VUB. Her academic background is mainly based on international and European human rights law, as she holds an LL.M. from the Department of International Studies at the Aristotle University of Thessaloniki. While studying for her master's degree, she spent a year at the Faculty of International Law at KU Leuven as an exchange student. She has also been a visiting researcher at the iCourts research team at the University of Copenhagen. Her academic research focuses on the "Divergencies of gender discrimination in AI".

Besides her academic interests, Anastasia is a digital rights activist, co-founding [DATAWO](#), a civil society organisation based in Greece for the advocacy of gender equality in the digital era. Anastasia Karagianni was a MozFest Ambassador for 2023 and Mozilla Awardee for the project "[A Feminist Dictionary in AI](#)" of the Trustworthy Artificial Intelligence working group.

The EU Artificial Intelligence Act through a Gender Lens

This paper provides a feminist analysis of the European Union's Artificial Intelligence Act (hereinafter the "AI Act"), assessing its capacity to address gender inequities and structural power imbalances in AI systems. Drawing on feminist theories, the paper evaluates the AI Act's limitations in mitigating gender biases that might disproportionately impact marginalised groups, particularly women of colour and women with disabilities. Through case studies in recruitment and employment, healthcare, border management control and risk assessment in domestic violence cases, the report highlights how AI applications can reinforce gender disparities. A detailed examination of specific provisions within the AI Act reveals critical gaps in addressing systemic discrimination and bias in AI governance. To promote a more equitable AI landscape, the report recommends integrating intersectional, feminist-informed revisions that prioritise interdisciplinarity, collective instead of individual approaches and strong oversight mechanisms that are oriented by human rights – including feminist – values. The proposed recommendations focus on strengthening the AI Act's framework to better safeguard marginalised groups and to ensure a regulatory approach that reflects the diverse experiences of all individuals.

Further information on the topic can be found here:

➤ fes.de