Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Gender in the Governance of Educational AI: A Critical Discourse Analysis of International Frameworks for AI Use in Teaching and Learning

Dr. Adnan Mohammed Bataineh
University of Buraimi
Kashif Ali Sabiri
kashif.sabiri@tuni.fi
Tampere University

Article

Keywords: critical discourse analysis, artificial intelligence, gender, education policy, governance frameworks, feminist discourse

Posted Date: November 16th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11384996/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Gender in the Governance of Educational AI: A Critical Discourse Analysis of International Frameworks for AI Use in Teaching and Learning

Abstract

International policy frameworks governing the use of artificial intelligence (AI) in education shape national policies and classroom practices. This research applies the three-dimensional Critical Discourse Analysis (CDA) model of Fairclough (2010) to three important international documents which are UNESCO Guidance for Generative AI in Education and Research (2023), European Commission Guidelines on the Ethical Use of AI (2025) and UNESCO Global Education Monitoring Gender Report: Technology on Her Terms (2024). This study investigates how gender is framed, which gender identities are centred or marginalized, how risks and opportunities are positioned in governance language, and what ideological assumptions underlie governance language through systematic analyses of lexical choices, transitivity structures, modality patterns, and ideological assumptions. It has been found that there is a significant discursive divergence across the data. Two of the three documents use strategies to make women and girls structurally invisible despite gender equity commitments. The third document employs feminist-structural framing that names certain gender identities and documents gender-specific harms. It also approaches gender inequality as a systemic condition requiring institutional accountability. According to the study, the gap between what is declared or professed and what is discursively enacted with regard to equity principles reflects a failure of governance with ramifications for policy. Suggestions are given for creating gender responsive frameworks for education governance for AI.

Keywords: *critical discourse analysis, artificial intelligence, gender, education policy, governance frameworks, feminist discourse*

Introduction

The Rise of AI in Teaching and Learning

Contemporary education has undergone a transformation due to the emergence of artificial intelligence (AI). The integration of AI into teaching and learning is picking up speed. Adaptive learning platforms, intelligent tutoring systems, and AI-generated instructional content are just some ways AI systems are being integrated into the classroom (Varsik & Vosberg, 2024). Governments, international organizations, and multilateral organizations are producing ethical and educational policy frameworks, advisory documents, and strategic recommendations to govern the responsible use of AI in education (UNESCO, 2023). These documents are highly authoritative at the institutional level, setting national policy, informing institutional decision-making and ultimately controlling how AI gets brought into classrooms globally. Although these frameworks have considerable power, the way they construct social categories gender in particular has, surprisingly, seldom been researched in a scholarly manner.

Gender and AI: A Persistent Blind Spot

The link between AI systems and gender inequality has received increasing attention across a range of fields. Studies show, repeatedly, that the AI tools replicate, if not exacerbate, existing gender biases contained in training data (UNESCO, 2024). Large language models (LLMs), for one, have been found to allocate male-gendered agents to professional, authoritative roles while assigning domestic, passive roles to female-gendered characters (Bearman et al., 2023). AI tools that tailor content to the intended audience can be helpful but also quite dangerous. For instance, a recent study analyzed the outputs of an elite science-generating AI and found that the tool coded explanations differently depending on the gender of the intended recipient (Hedlund, 2025). The results of our analysis indicate a structural issue. This issue does not simply lie in the technical

design of individual tools alone. It originates in the ideological assumptions within the institutions, policies, and governance that govern AI. As a result, gender is not simply a variable to be rectified in algorithms; rather, it is a discursive category produced through the institutional language of AI in education.

The Governance Gap

The literature has remained remarkably silent on how the international frameworks on AI-in-education construct gender despite their proliferation. Existing studies have analyzed AI discourse in higher education broadly (Bearman et al., 2023) and examined AI policy evolution in specific national contexts (Jeong, 2025) but none have analysed systematically a gender informed discourse on the corpus of international documents that governs AI use in teaching and learning. In addition, international documents such as the Guidance for Generative AI in Education and Research (2023) of UNESCO and the Varsik & Vosberg (2024) research papers on equity and inclusion in AI of the Organization for Economic Co-operation and Development (OECD) explicitly commit to gender equality. However, how this commitment is linguistically realized will be an empirical question. When international actors speak authoritatively about AI and education without a critical engagement with how their own language encodes gender assumptions, they risk institutionalizing the very inequalities they purport to address.

Purpose and Significance of the Study

This study attempts to fill the gap identified earlier in the paper. It carries out a CDA of international policies frameworks on the use of AI in teaching and learning, specifically focusing on how gender is constructed, positioned and possibly erased in the documents. The research makes use of Fairclough's (2010) three-dimensional model of analysis at the levels of the text, discursive practice and social ideology. This paper contributes in two ways. Theoretically, when

applying a gender-critical lens to the governance of educational AI, it aims to contribute to the CDA literature on a largely under-researched topic. Practically, it seeks to supply evidence-based recommendations for policymakers engaged in developing gender-sensitive governance of educational AI (Varsik & Vosberg 2024).

Research Questions

This study is guided by the following research questions:

1. What discursive strategies are used to frame gender in international AI-in-education policy documents?
2. Whose gender identities and concerns are centered, marginalized, or absent in these frameworks?
3. How do these documents position gender in relation to AI-driven risks and opportunities in teaching and learning?
4. What ideological assumptions about gender roles underpin the language of international AI education governance frameworks?

Literature Review

Gender Bias in AI Systems: Mechanisms and Manifestations

Scholarly research on gender bias in artificial intelligence has set a strong empirical foundation for the argument that artificial intelligence systems do not merely reflect societal inequalities but actually amplify them. A recent study by Kotek et al. (2023) found that large language models (LLMs) are three to six times more likely to select occupations that stereotype align with a person's assumed gender. The models' explanations for these choices, however, are factually inaccurate and amount to post-hoc rationalizations of structurally biased behavior. Building on this line of inquiry, Plaza-del-Arco et al. (2024) demonstrated that five state-of-the-art LLMs consistently attributed

gendered emotions in ways that reproduced well-established psychological stereotypes, associating anger with male-coded characters and sadness with female-coded characters across both open- and closed-source models. Consequently, these findings highlight that gender bias in AI is neither a superficial problem that can be rectified through manual interventions, nor is it merely a problem of bad data. Instead, these findings indicate that gender bias is a problem infused into AI at a much deeper level. Quite specifically, gender bias is embedded at the levels of semantic association, syntactic framing, or narrative framing. However, neither of these studies importantly tackles the institutional and discursive frameworks that license the deployment of such systems in specific social sectors including education. This is gap that drives the analytical focus of the present study.

Gender, AI, and the EdTech Landscape

The implementation of artificial intelligence (AI) in educational technology (EdTech) adds another dimension to the issue of gender bias. According to Peláez-Sánchez et al. (2023), implementation of equal access alone will not resolve gender inequity in EdTech for gender assumptions that are embedded. As if this were not bad enough, Otis et al. (2024) synthesized global evidence from 18 studies, which span more than 140,000 individuals and found that gender gaps in generative AI usage are nearly everywhere. Not only that, but these gaps also persist across regions, sectors and occupations even when access conditions are equalized. The possible underutilization of AI systems by women may create a vicious cycle: AI systems not developed with adequate consideration of women's needs and preferences will produce outputs that continuously marginalize women. Educational AI tools are already being used for teaching, assessment and feedback. This process may have serious implications for learning equity. However, the policy

document which govern the deployment of EdTech have been examined for whether they foresee, recognize or actively resist the reproduction of this cycle.

AI Governance in Education: Ideology, Power, and Missing Voices

Though there has been an increase in the scholarship on AI governance in education, the literature remains predominantly technical and procedural, viewing governance mainly as a question of regulation and institutional management, rather than as discourse and ideology. According to Filgueiras (2024), AI governance in education can be thought of from the prism of public policy. As AI systems are incorporated into educational systems, new configurations of bureaucratic power take an emergence with a tendency of concentration of decision-making in the hands of a technocratic elite away from teachers, students, and marginalized groups. Although this analysis highlights differences and imbalances of structural power, it does not extend to issues of how gender is constructed most likely neglected through these power configurations. The OECD Digital Education Outlook (OECD, 2023) similarly addressed the governance of generative AI in education from a comparative policy perspective, mapping how countries are developing digital education ecosystems and governance frameworks. However, where gender was mentioned at all, it was treated as one equity variable among many, with no critical scrutiny applied to its discursive construction. In the article, Marland & Arantes (2026) argue for a fundamental rethink of AI governance in education drawing on examples from global implementation contexts. Reframing this governance question, they conclude that current governance does not recognize the lived experiences of learners from marginalized communities. According to the analysis, it appears that the need for discourse-analytical approaches is warranted, which can show whose experiences and identities go visible or invisible in the policy language- precisely the methodological contribution the present study seeks to make.

The Intersection of Gender and AI Policy Discourse: What Is Missing

The most pertinent existing scholarship looks at the intersection of gender and AI at the level of policy and regulation. Karagianni (2025) performed a feminist legal reading of the EU AI Act. Although the regulation formally recognizes gendered risks, it adopts an individualized harm-prevention framework. This means that structural gender inequalities, including those reproduced through AI in education, remain intact. In an analysis of the EU AI regulatory frameworks, Baroni and Padovani (2025) came to similar conclusions: texts on governance endorse gender mainstreaming commitments but operate as mere rhetorical devices, entirely disconnected from enforcement guarantees. Cupac's (2025) most recent research, which examined the extent to which international AI governance frameworks address gender-related harms, uncovered continued inconsistencies across the documents, minimal references to intersectional gender identities, and almost no education-specific gender provisions. Additionally, none has employed Critical Discourse Analysis as their methodology. This suggests there is no evidence to indicate that the language mechanisms which enable gender construction, the ways in which the authors' choice of lexicon, grammar, pronoun and ideology are made.

Synthesis and Rationale for the Present Study

The literature presented so far presents a clear and consequential picture: While gender bias in AI is empirically well-established (Kotek et al, 2023; Plaza-del-Arco et al, 2024), gender inequity in EdTech adoption is structurally evidenced (Peláez-Sánchez et al, 2023; Otis et al, 2024), gender mainstreaming in AI governance is starting to be the subject of critical scrutiny (Baroni & Padovani, 2025; Karagianni, 2025; Cupac, 2025), no study so far has brought all these three strands together through a coherent methodological framework applied to international AI-in-education policy documents. The regulation of AI for teaching and learning constitutes a unique discursive

site where institutional authority, technological ideology, and educational values intersect and which is the site where gender assumptions may otherwise be institutionalized and naturalized at scale. This study seeks to bridge this gap by conducting a Faircloughian three-dimensional CDA on a purposively selected corpus of international AI-in-education frameworks to make visible the gendered discursive work that these authoritative texts do.

Methodology

Research Design

The present study employs a qualitative research design based on Critical Discourse Analysis (CDA) to explore the construction and representation of gender within international policy frameworks that govern the application of artificial intelligence (AI) in the domain of teaching and learning. According to Bearman et al. (2023), a qualitative approach is suitable in light of the fact that the research deals with the meanings, ideologies and power relations in policy language rather than numerical or statistical outcome. The use of CDA as the overall methodology is because it helps in systematic study of language not just as a communication tool but as a site where social inequalities such as gender inequalities are produced, reproduced and resisted (Simpson, 2025).

Theoretical Framework

The analysis adopts Fairclough's (2010) three-dimensional CDA model to carry out the analysis. The model has three levels of analysis namely: (1) textual level, which refers to the linguistic features of the text such as vocabulary, grammar, transitivity, and pronouns etc.; (2) discursive practice, which refers to the way a text is produced, distributed, and consumed in an institution; and (3) social practice, which refers to the discourses and ideologies that shape the text. This tripartite framework is fitting for policy document analysis as it enables the researcher to shift easily between close textual analysis and sociopolitical critique at the macro level. When the

framework is applied to the intersection of gender and AI education governance, it can highlight overt and covert gendered assumptions that are embedded in institutional language (Dahlborg et al., 2024; van Dijk, 2015).

Corpus: Data Sources and Selection Criteria

The analytical corpus comprises three publicly available policy and guidance documents produced by the leading international organizations that specifically address the use of AI in teaching and learning. Based on three criteria, documents were included in this study; (a) published between 2023 and 2025, (b) explicit references to the use of AI in education, and (c) major international governance bodies. The corpus under study is embodied in Table 1.

Table 1

Corpus of Policy Documents for Analysis

#	Document Title	Issuing Body	Year	Gender Stance
D1	Guidance for Generative AI in Education and Research	UNESCO	2023	Peripheral
D2	Guidelines on the Ethical Use of Artificial Intelligence (revised edition)	European Commission	2025	Incidental
D3	GEM Report 2024: Technology on Her Terms	UNESCO	2024	Central

Following the principles of purposive sampling that is standard in qualitative CDA research (Wodak & Meyer, 2015), three documents were selected on the basis of their institutional authority, direct relevance to AI use in teaching and learning, and most importantly, their contrasting stances towards gender, ranging from gender as a marginal equity variable to gender as the core organizing principle. The plausibility of such a corpus size includes the acknowledgment that depth as

opposed to breadth of analysis is a well-established norm in CDA policy studies (Hammersley & Atkinson, 2019).

The research questions most directly addressed by each document in the corpus has been mapped to the document to ensure systematic and transparent analytical coverage of all three texts. Document 1 (UNESCO, 2023) focuses on the governance of language needed around generative AI for educational purposes and the risks and opportunities for learners. Based on which it acts as the primary source for RQ1 and RQ3. Document 2 (European Commission, 2025) is the main source for RQ1 and RQ4. It is because the regulatory register and legally inflected language of document 2 provide the richest available data to examine discursive strategies and ideological assumptions about gender in AI Governance. The fundamental text for RQ2 and RQ3 is Document 3 (UNESCO, 2024) as it is the only document in the corpus for which gender is the organizing principle. Hence, it provides a good baseline for the comparative analysis of gender discourse in D1 and D2. All four research questions are addressed across the whole corpus, with none of them relying on one document alone, ensuring analytical triangulation and cross-document veracity.

Corpus–Research Question Alignment

All the documents are mapped to the research questions they most directly serve, ensuring systematized and transparent analytical coverage.

Table 3

Research Question–Corpus Alignment

Research Question	D1 — UNESCO (2023)	D2 — EC (2025)	D3 — UNESCO (2024)
RQ1 — Discursive strategies framing gender	☑ Primary	☑ Primary	☑ Supporting

RQ2 — Centered/marginalized gender identities	☑ Supporting	☑ Supporting	☑ Primary
RQ3 — Gender and AI risks/opportunities	☑ Primary	☑ Supporting	☑ Primary
RQ4 — Ideological assumptions about gender	☑ Supporting	☑ Primary	☑ Supporting

The holistic analysis of the three documents is ensured by mapping it to the four research questions in a way whereby each research question is answered by at least three documents.

Analytical Procedure

The analysis follows three successive steps following Fairclough's (2010) model of a three-dimensional analysis. The first phase (textual analysis) involves close reading of each document to uncover the mechanisms of gender construction used in the texts. Emphasis is on lexical selections, naming and referring strategies, uses of pronouns and transitivity patterns, which reveal the ways that gender identities and gender roles are discursively produced. It is important to note that the analysis not only illustrates who acts as a cause of harm or good, but it also shows those who remain "passive" receivers with important consequences for the institutional responsibility seen in governance documents. The following phase, the discursive practice analysis, analyses the extent to which 'gender framings' are foregrounded, marginalized or completely absent throughout the corpus (and repeatedly within it), through the intertextual comparison of the three documents. The third and last step social practice analysis involves interpreting the textual and discursive findings in light of more abstract, socio-political frameworks, informed by scholarship on the intersections of gender inequality, educational technologies and policy governance and AI design (Bearman et al., 2023; Jeong, 2025).

Trustworthiness

Two strategies are used to enhance credibility and dependability. Analytical transparency is ensured through the provision of thick descriptions of the analytical process and textual evidence for every interpretation. Second, I conducted a peer debriefing by presenting preliminary understanding to the colleague familiar with CDA and discuss any possible interpretive blind spots (Dahlborg et al., 20242).

Ethical Considerations

This article does not contain any studies with human participants performed by any of the authorsD2.

Findings

This section presents the results of the application of Critical Discourse Analysis (CDA) to the three international policy frameworks regulating the use of AI in education and research: D1 (UNESCO, 2023), D2 (European Commission, 2025) and D3 (UNESCO, 2024). The analysis of findings has been performed based on Fairclough (2010) three-dimensional model namely textual analysis, discursive practice and social practice. The presentation of the findings has been organized in the form of four research questions and supported through direct text from each document.

Finding 1: Discursive Strategies for Framing Gender

Across the three documents, three strategies become dominant: first, technical neutralization; second, juridical minimalism; and third, feminist-structural naming. Not all strategies are equally distributed. Almost exclusively the first two will be used by D1 and D2 while D3 uses latter.

1.1 Technical Neutralization (D1 — UNESCO, 2023)

In D1, gender gets absorbed into the wider technical category of 'bias', denuding it from its social specificity and repositioning it as a 'data quality problem'. D1 does name 'stereotypical gender

roles' once, as an example of biased output (section 2.7), but only as an illustration within a general warning to users and not as a structural analysis.

"*Build validation mechanisms to test whether GenAI systems used in education and research are free of biases, especially gender biases, and whether they are trained on data representative of diversity (in terms of gender, disability, social and economic status, ethnic and cultural background, and geographic location)*." (D1, p. 25)

This section is revealing in more than one way. First, the gender alluded to is acknowledged but is placed in a list of technical attributes representativeness of data, disability status, geography that are all put forward as conditions for validation of the model. An institutional validation mechanism and not an individual or group impacted by gender bias is the agent of change. The structurally absent position of the subject of women or girl is indicated by their disappearance as both named actors and subjects of redress. Fairclough (2010) refers to this process as 'nominalization,' which takes a complex social process (gender discrimination) and turns it into an abstract noun (bias), which conceals agency, cause and identity. A similar move occurs when D1 states:

"*Providers of GenAI must prohibit the design and deployment of GenAI systems that generate biased or discriminatory content based on race, nationality, gender or other protected characteristics.*" (D1, p. 22)

Here, gender is counted with race and nationality in a legal nondiscrimination formula. It is juridical listing: gender is noted, but not analytically prioritized or structurally explained.

1.2 Juridical Minimalism (D2 — European Commission, 2025)

D2 uses a blatant legalistic framing strategy. Gender is most conspicuous in the document's discussion of the Fairness and Non-Discrimination principle which asks.

"*Including the avoidance of unfair bias, regardless of age, gender, abilities, or any other historical disadvantaged attributes or characteristics.*" (D2, p. 24)

The term “regardless of age, gender, abilities” serves as a universal formula of neutralization: gender is included but is equivalently rendered to all other differences. 'Regardless of' is the standard non-discrimination formula and, read alone, it is protective: it says gender must not affect how a learner is treated. The argument here rests on what surrounds it. Gender sits in a list of equivalent attributes, and no group affected by gender-based unfairness is named anywhere in D2. Gender is therefore acknowledged as a principle but absent as a social subject. D2's guiding questions for teachers ask:

"*Do I notice any outputs from the AI tool that misrepresent, exclude, or negatively portray particular groups?*" (D2, p. 28)

In the teacher-facing questions, D2 leaves the noticing of gendered harm to an individual teacher, as a contingent and individualized observation rather than a structural condition. D2 does pose school-level questions in the same table, for example whether the school has ensured the tool was tested for bias across different learner groups. Those questions treat gender as one learner attribute among many and name no gendered group, so institutional responsibility is framed in generic terms while detection in classroom use rests with teachers. D2 has a wider ideological pattern whereby gender inequality is framed mainly as something a well-informed teacher might notice, and is not treated as a structural condition that an institutional actor must prevent.

1.3 Feminist-Structural Naming (D3 — UNESCO, 2024)

D3 Uses a completely different discourse strategy. As opposed to restricting gender to technical or legal categories, or neutralizing it, the text names gender as a structural social condition and makes

women and girls the explicit subjects of both risk and agency. The manuscript begins with a clear statement.

"*Technology is a tool. It promises easier access to education; personalized and enriched learning experiences. Yet, technology is not yet gender neutral.*" (D3, p. viii)

This phrase has ideological significance. In stating that technology is not yet gender neutral, D3 ‘says no' to the embedded assumption of technocratic neutrality found in D1 and D2. This is the taken-for-granted assumption that technology and policy language is without gender until proven otherwise. In D3, this view is inverted, meaning gender inequality is seen as the norm and equity must be constructed.

Finding 2: Centring, Marginalizing, and Erasing Gender Identities

2.1 Who Is Named — and Who Is Not

One of the study's most analytically interesting findings is about the visibility of certain gender identities across the three documents. The contrast is stark:

Table 2

Visibility of Gender Identity

Identity Group	D1 (UNESCO 2023)	D2 (EC 2025)	D3 (UNESCO 2024)
Women	Not named as group	Not named as group	Extensively named
Girls	Not named	Not named	Extensively named
Female teachers	Not named	Not named	Named explicitly
Non-binary students	Not named	Not named	Named explicitly
LGBTQ+ students	Not named	Not named	Named (cyberbullying context)

D1 or D2 doesn't specify women, girls or any particular gender identity as a subject for specific attention. D3 refers to women and girls on almost every page. On the matter of digital access, for instance, D3 states with accuracy.

"*There are 244 million fewer women than men using the Internet worldwide*." (D3, p. 2)

And on representation in AI specifically:

"*Women constitute only 26% of employees in data and artificial intelligence, 15% in engineering, and 12% in cloud computing across the world's leading economies.*" (D3, p. 3)

These statements are not rhetorical, which means to say these are not taken in vain. Rather, the statements themselves are quantified structural claims which are supported by data source citations. Furthermore, these statements are not subject to interpretation. They make material aspects of gender inequality visible in a manner that D1 and D2 entirely do not.

In the following sentence, D3 is the only document in the corpus that explicitly names non-binary students.

"*Cyberbullying is common, and more so for girls and non-binary students. A global poll of children aged 7 to 18 showed that a higher share of those of unspecified gender (50%) than girls (37%) and boys (29%) consider it very likely they will encounter online harm.*" (D3, Box 9, p. 3)

The naming of non-binary students supported by data involves a considerable deviation from the corpus-wide practice of gender binarism.

2.2 Female Teachers: Named and Erased

It is particularly significant how female teachers are treated. D2, a document specifically developed to assist teachers in the ethical application of AI, positions the teacher as a generic professional throughout. The document speaks generically of teachers. It does not anywhere particularize gendered experiences within the profession. In contrast, D3 recognizes.

"*Female teachers also had to take on most of the household and child-rearing duties in addition to their teaching duties.*" (D3, p. 29)

The difference matters analytically: D2 gives guidance on AI governance for teachers as the main audience with no mention of the gendered conditions under which most of this audience, a majority of whom are globally women, are in paid work and live.

Finding 3: Gender as Risk and Opportunity

3.1 The Risk Landscape

The three documents differ fundamentally in how they frame gender-related risk. D1 acknowledges that "*GPT models can further marginalize already disadvantaged people*" (D1, p. 17) and stops short of naming women and girls as specifically at risk. The risks are framed as universal, affecting marginalized "communities" abstractly.

D2 introduces the concept of deepfakes in its glossary, defining them as:

"*AI-generated or manipulated images, audio or video content that resembles existing persons, objects, places, entities or events and would falsely appear to a person to be authentic or truthful.*" (D2, Glossary, p. 40)

Notably, the definition is framed entirely in gender-neutral terms or is silent on the disproportionate targeting of women and girls in the context of non-consensual deepfake pornography, despite ample international evidence.

D3 names the same phenomenon with precision and with gender:

"*The danger of deepfakes made by artificial intelligence is spilling over into school environments. In Brazil and Spain, for instance, schoolgirls have been victims of deepfake pornographic images circulating in schools, while female teachers have been recorded as victims of the abuse in the United States.*" (D3, p. 3)

The distinction is not one of emphasis only but rather a discursive choice with implications for policy. The gender-neutral framing of deepfake risk by D2 means that any policy response to emerge from that framework does not have an institutional basis to prioritize the protection of girls and women. D3's name helps establish just that.

3.2 The Social Media–Well-being Nexus

D3 provides rich data on the gendered connection between social media and well-being and academic performance.

"*TikTok's algorithm targets teenagers with body image and mental health content every 39 seconds, and with content related to eating disorders every 8 minutes.*" (D3, p. 36)

"*Facebook's own research found that 32% of teenage girls said that when they felt bad about their bodies, Instagram made them feel worse.*" (D3, p. 2)

"*Girls are twice as likely as boys to suffer an eating disorder, which is exacerbated by the use of social media.*" (D3, p. 2)

These findings are not present in D1 or D2 in any form. The fact that they are missing is significant in itself: a governance framework for AI in education that does not address harm such as the algorithmically-driven social media on the mental health of girls cannot adequately protect them.

3.3 Gender as Opportunity: STEM and AI Leadership

D3 perceives gender equity in a unique manner. It does not perceive it solely as a risk-mitigation challenge but also as a structural opportunity. It contends.

"*Doubling the share of women in the tech workforce by 2027 could boost GDP by EUR 600 billion.*" (D3, p. 3)

And further:

"*Promote female leadership in artificial intelligence and technology development in order to assure gender-sensitive digital transformation and address gender stereotypes in algorithms.*" (D3, p. 51)

D1 touches on this theme obliquely, urging stakeholders to "promote gender equality in developing advanced AI competencies and create a gender-balanced pool of professionals" (D1, p. 26) though the proposal is one of five bullet points under a single competency-development heading (section 4.4), and is not accompanied by any supporting data or structural rationale. D2 Make no mention of gender and STEM representation.

Finding 4: Ideological Assumptions About Gender Roles

4.1 The Ideology of Technocratic Neutrality

Both D1 and D2 are anchored in what is conceptualized in this study as the ideology of technocratic neutrality that is, the implicit assumption that AI systems and governance frameworks are gender-neutral unless a particular bias is detected and ‘fixed'. The theory presents gender inequalities as a possibility rather than a manifestation of social structure. D2, an educator may "notice" prejudice; in D1, a legitimacy mechanism may "test" for it. It also tends to place the work of detecting gender unfairness with individual teachers, although D2 does pose some school-level questions. It is not clear who is harmed when these processes fail.

4.2 The Ideology of Feminist Structuralism

D3's analysis rests on a contradistinctive feminist-structural analysis. The framework considers gender inequality as a historically produced condition that is systemically produced through social institutions, including AI systems and education policy. This is stated explicitly in the document.

"*Biased social and cultural norms inhibit equitable access to and engagement with technology in and outside of school, with girls always left on the wrong side of the divide.*" (D3, p. v)

The phrase '*always left on the wrong side*' is the most ideologically loaded of all in a productive sense; it negates the conditionality found in D1 and D2, asserting gender inequality as a structural constant that policies must work against rather than monitor.

4.3 The Feedback Loop: A Structural Argument

D3 also introduces what it calls a "feedback loop," which may be the most powerful definitional analytic in the entire corpus.

"*All these factors create a feedback loop: where girls are exposed to negative gender norms, steered away from studying STEM subjects, and deprived of the opportunity to shape the very tools that expose them to these stereotypes.*" (D3, p. viii)

AI tools embed gender stereotypes which steer girls away from STEM which creates an underrepresentation of women in AI development which produces AI tools that embed gender stereotypes. D1 and D2 do not approach this question in terms of his structural critique of education-technology-gender relation. The fact that D1 and D2, which are documents with a clear commitment to equity, do not mention ideology represents a major gap in AI in education governance.

Discussion

Governance Language as Ideological Practice

The results confirm that gender is not only represented but also actively constructed in international AI-in-education policy through particular discursive choices. The three documents which were analysed show the same pattern; D1 and D2 show a pattern of discursive erasure. It means that there is recognition of gender at the level of vocabulary (bias, fairness, non-discrimination). However, there is a stripping out of structural content. It also includes any named subjects and material consequences. According to Fairclough (2010), the ideological work in

institutional texts does not take place through exclusion, but rather the normalization of some ways of seeing. In this case, gender-neutral governance language serves as a certain sign of impartiality, rather than a form of concealment.

The results and other literature show that artificial intelligence systems are not socially neutral (Noble, 2018; Buolamwini and Gebru, 2018). When governance frameworks reproduce the assumption of neutrality at the level of language, they enhance the gender-biased outcomes that their stated principles would want to prevent. The disparity between the equity commitments outlined in D1 and D2, and the linguistic practices via which those documents are drafted, is not incidental: it is ideological.

The Cost of Naming No One

One of the more noteworthy findings is that in the main texts of either D1 or D2, there are no references to women or girls as specific social subjects. This lack is not innocent. The framing of a problem determines what solutions are considered thinkable (Bacchi, 2025) in other words, the framing establishes the context of the problem. A governance framework that frames gender bias as a data quality problem (D1) or an individual teacher's perceptual task (D2) forecloses structural policy responses, e.g. investment in female STEM representation, gendered risk assessments for AI tools, institutional accountability for algorithmic harm to girls.

D3 demonstrates that alternative framings are possible and productive within the same institutional genre. By naming women and girls explicitly, documenting their specific vulnerabilities (deepfakes, social media harm, mathematics anxiety), and positioning them as active agents in technological transformation, D3 creates the discursive conditions for genuinely gender-responsive policy. The fact that two documents produced by the same parent organization, UNESCO, in less than a year of each other are so different from each other is a finding worth

bringing to the fore: institutional commitment to gender equity does not automatically imply gender-equitable discourse.

Intersectionality and the Limits of Binary Framing

A second noteworthy finding is the almost-complete lack of intersectionality analysis in D1 and D2 and the partial presence in D3. The act of D3 naming non-binary students in the context of being at risk for cyberbullying shows a considerable departure from the gender binarism of the corpus overall. Nevertheless, D3 never uses the term 'intersectional', and this study reads its treatment as selectively intersectional: for instance, it considers gender alongside geography, poverty and crisis contexts, but does not address transgender identities in a systematic way, nor the compounding of gender and disability in AI-mediated educational contexts. This indicates that even the most gender-progressive text in this corpus operates with ideological restrictions - and indicative of a fruitful agenda of CDA work in this domain in the future.

Implications for Policy and Practice

The findings have implications for the development and revision of AI-in-education governance. The gender equity commitments must be operationalized discursively, by enacting naming, evidencing, and structurally explaining gender inequality not only referred to in principles clauses. The second feedback loop highlighted in D3 is of girls experiencing gender stereotypes, being excluded from STEM, and therefore not being able to shape AI systems that reinforce these stereotypes. This should serve as a structuring logic for future policy documents and not a simple footnote. Governance frameworks must go beyond the individual teacher as the locus of detection of gender bias and institutional accountability mechanisms for gendered AI harms in education should be established.

Conclusion

The purpose of the study was to investigate how gender is discursively constructed in three international policies on AI in education or use of AI in education: 2023 UNESCO Guidance for Generative AI in Education and Research; 2025 European Commission Guidelines on Ethical Use of AI; and 2024 UNESCO Global Education Monitoring Gender Report. By drawing on Fairclough's (2010) three-dimensional CDA model, the analysis demonstrates a stark and consequential divergence across the corpus.

The key finding is that institutional commitment to gender equality does not necessarily lead to gender-equal discourse. The technical neutralization and juristic minimalism strategies in D1 and D2 render women and girls invisible as named social subjects. These two policy documents invoke gender equality as a governing value. Although gender is recognized, it is structurally deleted. In contrast, D3 utilizes feminist-structural framing approaches that identify specific gender identities, substantiate gender-specific harms, and opportunities, and find the causes of inequality in social and algorithmic structures, not in individual actions or technical failures.

This difference in attitude is not unimportant. The governance frameworks are generative: the language in which they are framed shapes which problems get visibility, whose experiences get protection, and which policy responses become conceivable. A framework that puts gender in a list of biases cannot create the structural accountability mechanisms to protect girls from deep fake harm, to reduce algorithmic amplification of gender stereotypes and to close the gender gap in AI leadership. Hence, narrowing the gulf between expressed gender equity pledges and the rhetoric through which they are operationalized is not a matter of stylistic finesse; it is a condition for effective, equitable AI governance in education. The analysis used here should be applied to national-level research on AI education policy, and attend to the intersectional effects of gender, race, disability, and class in the discourse on AI governance.

Data Availability

The data analysed in this study are the full texts of three publicly available policy documents: UNESCO (2023), Guidance for generative AI in education and research (https://doi.org/10.54675/EWZM9535); the European Commission (2025), Guidelines on the ethical use of artificial intelligence, revised edition (https://doi.org/10.2766/7967834); and UNESCO (2024), Global Education Monitoring Report: Gender report, Technology on her terms (https://doi.org/10.54676/WVCF2762). All three are open access.

The following materials have been shared as supplementary files for peer review: (a) the data package (04_Data_Package.zip), containing the raw master text file of the three documents, cleaned text files for each document, and the Python script used for term counts, the concordance and quotation checks; (b) a Supplement (02_Data_and_Methods_Supplement.docx) describing the corpus, data preparation, variable definitions, coding scheme, analytic procedure and limitations; and (c) an evidence workbook (03_Evidence_Register_and_Verification_Tables.xlsx) containing the codebook, a quotation register checked against the source texts, term counts, and a concordance of "gender" in D1 and D2.

https://drive.google.com/drive/folders/1CjKzjrc82ggtRJ-JjqjT-RXXwYjkqfww?usp=sharing

There are no restrictions on sharing these materials. The source documents are published under open licences (CC BY-SA 3.0 IGO for the UNESCO documents and CC BY 4.0 for the European Commission document). The study involved no human participants, so there are no participant data. The interpretive coding was done by close reading and is not reproduced by the script.

References

Bacchi, C. (2025). What's the Problem Represented to Be? A New Thinking Paradigm. Routledge. https://doi.org/10.4324/9781032678382

Baroni, A., & Padovani, C. (2025). AI, democracy and gender equality: EU regulatory frameworks and the wager of gender mainstreaming. European Journal of Communication, 40(4), 312–328. https://doi.org/10.1177/02673231251349323

Bearman, M., Ryan, J., & Ajjawi, R. (2023). Discourses of artificial intelligence in higher education: A critical literature review. Higher Education, 86(2), 369–385. https://doi.org/10.1007/s10734-022-00937-2

Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. Proceedings of Machine Learning Research, 81, 1–15. https://proceedings.mlr.press/v81/buolamwini18a.html

Cupac, J. (2025). The gender code: Gendering the global governance of artificial intelligence. arXiv. https://doi.org/10.48550/arXiv.2512.09570

Dahlborg E, Boman Å, Eriksson H, & Tengelin E (2024). Encircling discourses: A guide to critical discourse analysis. (2022). Scandinavian Journal of Caring Sciences, 37(1), 286–295. https://doi.org/10.1111/scs.13194

European Commission. (2025). Guidelines on the ethical use of artificial intelligence (Revised ed.). Publications Office of the European Union. https://doi.org/10.2766/7967834 (PDF ISBN: 978-92-68-33189-7)

Fairclough, N. (2010). Critical discourse analysis: The critical study of language (2nd ed.). Routledge. https://doi.org/10.4324/9781315834368

Filgueiras, F. (2024). Artificial intelligence and education governance. Education, Citizenship and Social Justice, 19(3), 349-361. https://doi.org/10.1177/17461979231160674

Hammersley, M., & Atkinson, P. (2019). Ethnography: Principles in practice (4th ed.). Routledge.

Hedlund, A. (2025). Gender equity in GenAI science explanations. School Science Review, 395, 39–46. https://www.ase.org.uk/system/files/SSR%20395%20p39-46%20Hedlund.pdf

Jeong, H. (2025). From curriculum to strategy: tracing shifts in AI Education policy in South Korea through keyword and topic analysis. Asia Pacific Journal of Education, 1–30. https://doi.org/10.1080/02188791.2025.2580988

Karagianni, A. (2025). Gender in a stereo-(gender)typical EU AI law: A feminist reading of the AI Act. Cambridge Forum on AI: Law and Governance, 1, e5. https://doi.org/10.1017/cfl.2025.12

Kotek, H., Dockum, R., & Sun, D. (2023). Gender bias and stereotypes in large language models. In Proceedings of The ACM Collective Intelligence Conference (pp. 12–24). ACM. https://doi.org/10.1145/3582269.3615599

Marland, B. & Arantes, J. (2026). Reframing AI governance in education: Insights from global implementation contexts. Learning, Media and Technology. https://doi.org/10.1080/17439884.2025.2595443

Noble, S. U. (2018). Algorithms of oppression: How search engines reinforce racism. New York University Press.

OECD. (2023). OECD digital education outlook 2023: Emerging governance of generative AI in education. OECD Publishing. https://doi.org/10.1787/c74f03de-en

Otis, N. G., Delecourt, S., Cranney, K., & Koning, R. (2024). Global evidence on gender gaps and generative AI (HBS Working Paper 25-023). Harvard Business School. https://www.hks.harvard.edu/centers/cid/publications/global-evidence-gender-gaps-and-generative-ai

590 Peláez-Sánchez, I. C., George-Reyes, C. E., & Glasserman-Morales, L. D. (2023). Gender digital
591 divide in education 4.0: A systematic literature review of factors and strategies for inclusion. Future
592 in Educational Research, 1(2), 89–114. https://doi.org/10.1002/fer3.16
593 Plaza-del-Arco, F. M., Curry, A. C., Abercrombie, G., & Rieser, V. (2024). Angry men, sad women:
594 Large language models reflect gendered stereotypes in emotion attribution. In Proceedings of the
595 62nd Annual Meeting of the Association for Computational Linguistics (pp. 7682–7696). ACL.
596 https://doi.org/10.18653/v1/2024.acl-long.415
597 Simpson, N. H. (2025). Framing AI in higher education: A critical discourse analysis of inclusivity
598 and power in university AI guidance documents. Educational Linguistics.
599 https://doi.org/10.1515/eduling-2024-0007
600 UNESCO. (2021). Recommendation on the ethics of artificial intelligence. UNESCO.
601 https://unesdoc.unesco.org/ark:/48223/pf0000381137
602 UNESCO. (2023). Guidance for generative AI in education and research. UNESCO.
603 https://doi.org/10.54675/EWZM9535
604 UNESCO. (2024). Generative AI and the future of education: Alarming evidence of regressive
605 gender stereotypes. UNESCO. https://www.unesco.org/en/articles/generative-ai-unesco-study-
606 reveals-alarming-evidence-regressive-gender-stereotypes
607 UNESCO. (2024). Global Education Monitoring Report gender report: Technology on her terms.
608 UNESCO. https://doi.org/10.54676/WVCF2762
609 van Dijk, T. A. (2015). Critical discourse analysis. In D. Tannen, H. E. Hamilton, & D. Schiffrin
610 (Eds.), The handbook of discourse analysis (2nd ed., pp. 466–485). Wiley-Blackwell.
611 https://doi.org/10.1002/9781118584194.ch22

Varsik, S., & Vosberg, L. (2024). Supporting equity and inclusion with AI in education: Opportunities and challenges (OECD Education Working Paper No. 310). OECD. https://doi.org/10.1787/3bb4f1ad-en

Wodak, R., & Meyer, M. (Eds.). (2015). Methods of critical discourse analysis (3rd ed.). SAGE.