**Input from the CSTD Gender Advisory Board**

# Data Symmetries: procedural, representational, distributive, temporal

This piece was authored on behalf of the CSTD Gender Advisory Board by Gender Advisory Board members Caitlin Kraft-Buchman, Co-Chair CSTD Gender Advisory Board, CEO / Founder of Women At The Table, and Ritu Agarwal, Wm. Polk Carey Distinguished Professor of Information Systems and Health at the Johns Hopkins Carey Business School and founding co-director of Center for Digital Health and Artificial Intelligence (CDHAI) with substantial input from Li Ning, Director, Division of Asian and African Affairs, China Science and Technology Exchange Center (CSTEC), Ministry of Science and Technology of the People's Republic of China

## 1. Purpose

This report consolidates various components from fundamental principles, inclusive standard setting,upstream data, and disaggregated data under a single operational concept: **data symmetry.**

The report registers a finding from the recent scientific literature which changes the timeline on which the underlying problem must be addressed. Under current AI training practices, asymmetries in data compound across model generations. The data gap is not a static deficit awaiting gradual correction. It is a closing window.

We offer this report in support of commitments Member States have already made. The Global Digital Compact (General Assembly resolution 79/1, annex I) establishes gender equality and the empowerment of all women and girls as a cross-cutting principle of digital cooperation and sets, as its fourth objective, the advancement of responsible, equitable and interoperable data governance. Under that objective States acknowledged that data divides, including gender data gaps, produce inequitable distribution of benefits and biased results, and committed to develop data and metadata standards designed to prevent and address bias throughout the data life cycle, including through regular data auditing. The mandate of the Working Group on Data Governance, set out in the same objective, extends expressly to considerations of sharing the benefits of data. The WSIS+20 outcome document (General Assembly resolution 80/173) reaffirms that the full, equal, safe and meaningful participation of women in the digital space is essential to closing the gender digital divide, commits to catalysing women's economic agency, and calls for the further development of internationally agreed targets, indicators and metrics for digital development. What follows is therefore not a new normative claim. It is the measurement and implementation architecture without which the existing commitments cannot be verified, the scaffolding without which the commitments cannot be certified as fulfilled, and which cannot be built later if it is not begun now.

## 2. The empirical baseline

The asymmetry is documented, not asserted. Women average 37% of participants in randomised clinical trials, and roughly three quarters of studies report no sex-stratified outcomes; the clinical record on which health AI trains describes a male default body. The consequences are measurable at the output layer: clinical AI trained on sex-misrepresentative data has been found to reduce diagnostic accuracy for women by 11.3 percentage points (Jabour, 2023), and large language models return different clinical assessments from identical case notes depending only on the patient's sex label. In the economic domain, the ILO values unpaid care work, roughly three quarters of it performed by women, at approximately 16.4 billion hours per day, on the order of US$11 trillion annually. Almost none of this activity appears in the statistical record from which financial and planning models learn. In language, the disparity between speaker populations and web representation means women in low-resource language communities are absent twice over: by language, and by gender within the language data that exists.

In the economic and credit domains, systemic data asymmetries are similarly codified into automated scoring and resource allocation systems, creating structural disadvantages at the point of evaluation. Personal financial models, increasingly driven by algorithmic credit scoring, reproduce historical gender inequalities under the guise of mathematical objectivity; in East Africa, for instance, a widely deployed digital lending algorithm in Kenya was documented to offer female applicants credit limits that were over a third, greater than 33%, lower than those offered to male applicants, despite the women demonstrating superior loan repayment performance (Canavesio, 2026). At the gateway of the formal labor market, automated hiring systems reinforce these allocative disparities by penalizing non-linear career trajectories and female-coded credentials. Empirical testing of resume-screening algorithms reveals a systematic preference for female-associated names in only 11.1 per cent,male-associated names in 51.9% of evaluations and white-associated names in 85.1% of cases, actively penalizing caregiving-related career gaps and docking scores for female-typed extracurricular markers, such as listing 'softball' rather than 'baseball' (University of Washington, 2024). These discriminatory outcomes are further entrenched by the introduction of gamified and standardized AI 'fit' assessments, which, as documented in field research, hard-code uniform thresholds that actively override the context-sensitive, human managerial discretion required to adapt evaluation criteria to localized labor markets and diverse applicant needs (Harvard Business Review, 2023).

Beyond these allocative consequences, representational and linguistic asymmetries in large language models (LLMs) demonstrate how deficits in upstream training data compound into biased model outputs and deployment choices. At the data layer, major LLMs reflect decades of unequal textual representation by associating women with 'home' and 'family' four times more frequently than men in model output (UNESCO, 2024). At the algorithmic output layer, these representational gaps actively reinforce male identity as the default benchmark of authority; interaction audits of conversational models demonstrate that the default holder of power is frequently imagined as male. When queried regarding a generic political figure (such as how a mayor’s son should address the mayor), LLMs demonstrably default to paternal titles like 'Dad' or 'Father' and explicitly rationalize that female political authority represents an 'exceptional case' requiring unique sentence constraints or explicit linguistic annotations (Chen, 2026). This

pattern is mirrored in pronoun assignment, where models inject gender bias into ambiguous scenarios without user prompting, associating prestigious professional roles (such as 'physicist') with male pronouns and supporting roles (such as 'secretary') with female pronouns (Ho et al., 2025). Finally, at the downstream deployment layer, this representational asymmetry is further physicalized in product design, where companies deploy submissive, compliant female identities for default virtual voice assistants while reserving authoritative, decision-making tasks for male-coded systems—a commercial design choice that commodifies female digital labor and solidifies patriarchal expectations of compliance among users (West et al., 2019; Browne et al., 2023).

Where the deficit concerns the body, the operative variable is sex; where it concerns labour, roles and participation, it is gender. Both are in play across the record.

These are not gaps at the periphery of the data economy. They concern half of every population that data-driven systems in health, credit, social protection and justice now serve.

**3. The concept: four symmetries**

**Data symmetry names a relation, not an aspiration.** A data system is symmetric with respect to a population when that population's presence in the record, its standing in the decisions that govern the record, its share of the value the record generates, and the durability of all three across successive uses are equal to those of any other population the system serves. Whereas data justice describes an ethical horizon and data equity a distributional goal, and data fairness is a property of outputs tested after deployment, symmetry specifies a measurable and quantifiable relationship between populations at four identifiable points in the data life cycle. The distinction is operational rather than terminological. Fairness asks whether a deployed system treats people equally, and can be satisfied by correction at the output layer. Symmetry asks whether the record and the foundational infrastructure – both system-related and policy-related – around the record were built such that equal treatment is achievable at all. A system can be adjusted to appear fair on a record that remains asymmetric. It cannot be made valid on one.

***Procedural symmetry.*** The populations a data system describes hold authority over its governance, not only their presence within it. The condition concerns who defines the problem a dataset is built to answer, who sets the categories it records, who determines what counts as adequate coverage, and who has standing to contest a system in deployment. Presence and agency are not interchangeable: a consultative seat in a standard-setting body confers voice without conferring power to determine an outcome, and the distinction is decisive for developing economies, which participate in technical standard-setting as rule-takers while the specifications that will govern their data are settled elsewhere.  Inclusive standard-setting supplies the institutional site for this condition. What is missing is the recognition that it sits upstream of the others: the room that defines adequate coverage, sets metadata requirements and writes terms of use is the room in which representation and benefit are decided, before any data is collected.

***Representational symmetry.*** The populations a system will affect are present in the data that builds and evaluates it, at levels sufficient for equal performance. This is a condition of scientific

validity, not a courtesy. A model trained on a record that documents one sex is not a general model with a fairness deficit; it is a model of that sex, misspecified.¨eferences to disaggregated data supply the measurement basis for this condition; what is missing is its recognition as a validity requirement rather than a desirable feature.

***Distributive symmetry.*** Value returns to those who generate data. We observe that returns concentrate downstream while upstream contributors receive limited or indirect benefit, compounding adverse consequences for digitally invisible populations who are non-participants in the data generation and curation activity. Distributive symmetry converts that observation into a design requirement: guaranteed payment to data collectors and contributors at the point of work, plus a proportional share of downstream commercial value. This is no longer a novel legal claim. Under the Convention on Biological Diversity, COP16 established the Cali Fund (2024, launched 2025), under which commercial users of digital sequence information are expected to contribute to a multilateral fund reserving at least half its proceeds for the indigenous peoples and local communities who steward the underlying resource. The Nagoya Protocol supplies the treaty foundation. The principle that commercial use of a digitized, community-derived resource triggers an enforceable benefit-sharing obligation has been agreed by 196 governments for genetic data. Its extension to human-generated data is a matter of application, not invention.

***Temporal symmetry.*** The record improves, or at minimum does not degrade, across successive uses. Temporal symmetry requires intentional and vigilant attention from frontier model inception to every new release to measure the improvement or degradation. This condition is new to the governance discussion and calls for urgency in data inspection and quantification of the harm.

**4. Why the problem compounds**

Peer-reviewed research demonstrates that when models train on data generated by earlier models, degradation follows a predictable sequence: the sparse regions of the data distribution are lost first, while embedded biases amplify across generations (Shumailov et al., Nature, 2024, and subsequent literature). Fluency is preserved while rare and underrepresented patterns quietly vanish.

The claim should be stated at its evidenced strength. Collapse dynamics are demonstrated most cleanly under pure or near-pure recursion; production pipelines still mix fresh human data with synthetic, and the rate of degradation in deployed systems is not yet established. What the literature does establish is the direction and the asymmetry: recursion thins sparse regions rather than filling them, and the identified remedy, continued infusion of fresh, provenanced human data, works only where such data exists. For women's health beyond the male clinical default, women's informal and unpaid economic activity, and women's voices in low-resource languages, it largely does not exist in machine-usable form. It was never written down. Each training generation built on the thinned record raises the cost of correction, until correction becomes a re-foundation problem rather than a data problem. A related concern is the

compounding of bias as humans interact with biased AI systems (Glickman and Sharot, 2025), learning the “built-in” bias and incorporating it into their decisions – setting in motion a vicious feedback loop when these decisions become the training data for the next model iteration.

Two measurement gaps follow, and both are governance gaps. No institution has ever measured what share of a frontier training corpus that documents female physiology or records women's economic activity, because no such audit has been performed; the baseline number does not exist. And no institution measures how representational coverage moves across successive model generations and the extent of harm amplification that occurs when human agents internalize AI bias. Action is guided by data and becomes increasingly inequitable with every interaction.  Any assurance that current systems are safe to deploy for the populations in question rests on evidence that has not been gathered.

An acute manifestation of these systemic asymmetries appears in the digital safety sphere, where the asymmetric burden of AI-facilitated violence, image-based abuse, and targeted disinformation is weaponized to silence women in public life. The rapid consumerization of deepfake technology has created a highly gendered vehicle for digital abuse; non-consensual deepfake pornography now accounts for a striking 98% of all deepfake videos online, with 99% of all targeted individuals being women (Bentzen & EPRS, 2026). This technological weaponization is disproportionately directed at women occupying public and professional spaces, with nearly 25% of surveyed female journalists, activists, and human rights defenders reporting exposure to AI-assisted online violence and non-consensual image manipulation (UN Women, 2026). Furthermore, the geopolitical utility of these digital harms is increasingly recognized by authoritarian regimes, which systematically weaponize gendered AI-driven disinformation campaigns and synthetic media to degrade, humiliate, and discredit independent female political leadership, thereby enforcing digital transnational repression and driving women out of democratic debate entirely (Chenoweth & Marks, 2022; Bentzen & EPRS, 2026). From a data governance perspective, mitigating these harms requires establishing mandatory provenance documentation, robust digital watermarking, and clear traceability standards to distinguish human-generated media from synthetic outputs in public and political discourse.

## 5. Recommendations

The Gender Advisory Board recommends that Member States and United Nations bodies adopt data symmetry as an organising concept for data governance in development, and act on each of its four dimensions as follows.

1. ***Who decides (procedural symmetry)****.* Developers of frontier and large-scale AI systems should be required, not invited, to report what their training data contains, including its demographic coverage and the share that is synthetic, and how that coverage changes between model generations. Member States and United Nations bodies should agree on a common reporting format and a forum in which this reporting takes place. Ideally, the communities being reported on, including women's rights

organisations, should have decision-making authority in agreeing that format, not a consultative role alone.

2. ***Who is in the data (representational symmetry)***. AI systems used in public services should not be deployed in high-stakes settings until they meet minimum performance standards for women and for each population they serve. A system that performs well for one group and poorly for another is not a valid system, which makes representativeness a scientific requirement rather than an ethical preference. Datasets should record which populations they cover, in a standard form that can be checked across systems, and that record should be independently verified rather than self-reported.
3. ***Who benefits (distributive symmetry)***. Community organisations, including women's rights organisations, should be trained, certified and paid to collect and hold data from their own communities. They should be paid at the time of collection, retain ownership through trusts or cooperatives governed by the CARE principles, and receive a share of revenue when the data is used commercially. The Cali Fund, agreed at CBD COP16 under the Nagoya Protocol, already applies this model to genetic data.
4. ***Whether it holds over time (temporal symmetry)***. How well AI training data represents different populations should be measured independently as a public statistical function, beginning with a first baseline, which does not yet exist, and repeated with each new generation of models. AI systems used for public purposes should disclose which of their training data is human-generated and which is synthetic, and this information should be preserved when data moves across borders.

***Priorities***. International cooperation should fund the foundations first: data commons and the infrastructure to collect new data. Performance standards and ongoing measurement are permanent safeguards; post-training fixes are temporary. The foundations must be built now, while the gaps can still be closed by collecting data rather than by rebuilding models.

**Bibliography**


Abuwatfa, W., Zamel, N., & Al-Othman, A. (2021). Lessons learned from the underrepresentation of women in STEM: AI-enabled solutions and more. *Energy and AI*, 5, Article 100086. [4] https://doi.org/10.1016/j.egyai.2021.100086

Agnafors, S., Bladh, M., Svedin, C. G., & Sydsjö, G. (2019). Mental health in young mothers, single mothers and their children. *BMC Psychiatry*, 19, 112. [5] https://doi.org/10.1186/s12888-019-2082-y

Alozie, N. O., & Akpan-Obong, P. (2017). The digital gender divide: Confronting obstacles to women's development in Africa. *Development Policy Review*, 35(2), 137–160. [4] https://doi.org/10.1111/dpr.12204

Anderson, R. C., & Klofstad, C. A. (2012). Preference for leaders with masculine voices holds in the case of feminine leadership roles. *PLoS ONE*, 7(12), Article e51216. [5] https://doi.org/10.1371/journal.pone.0051216

Balsmeier, B., & Woerter, M. (2019). Is this time different? How digitalization influences job creation and destruction. *Research Policy*, 48(8), Article 103765. [5] https://doi.org/10.1016/j.respol.2019.03.010

Bitton, D. B., Hoffmann, C. P., & Godulla, A. (2025). Deepfakes in the context of AI inequalities: Analysing disparities in knowledge and attitudes. *Information, Communication & Society*, 28(2), 295–315. [6] https://doi.org/10.1080/1369118X.2024.2420037

Brandão, M., Mendes, F., Martins, M., Cardoso, P., Macedo, G., Mascarenhas, T., & Mascarenhas Saraiva, M. (2024). Revolutionizing women's health: A comprehensive review of artificial intelligence advancements in gynecology. *Journal of Clinical Medicine*, 13(4), Article 1061. [6] https://doi.org/10.3390/jcm13041061

Brown, J., Park, T., Chang, J., Andrus, M., Xiang, A., & Custis, C. (2022). Attrition of workers with minoritized identities on AI teams. In *Proceedings of the EAAMO '22* (pp. 1–9). Association for Computing Machinery. [7] https://doi.org/10.1145/3551624.3555304

Browne, J., Cave, S., Drage, E., & McInerney, K. (Eds.). (2023). *Feminist AI: Critical perspectives on algorithms, data, and intelligent machines*. Oxford University Press. [7]

Canavesio, E. (2026, April 30). Overcoming the algorithmic gender bias in AI-driven personal finance. *The Conversation*. https://theconversation.com/overcoming-the-algorithmic-gender-bias-in-ai-driven-personal-finance-281250

Chen, Z. (2026). *Future Faces of Research: Gender Bias in AI*. Feminist Research Institute, University of California, Davis. https://feministresearch.ucdavis.edu/gender-bias-ai-2026

Chenoweth, E., & Marks, Z. (2022). Revenge of the patriarchs: Why autocrats fear women. *Foreign Affairs*, 101(2), 114–128.

Coeckelbergh, M. (2020). *AI ethics*. The MIT Press. [8]

Danielescu, A. (2020). Eschewing gender stereotypes in voice assistants to promote inclusion. In *Proceedings of the 2nd Conference on Conversational User Interfaces*. Association for Computing Machinery. [8] https://doi.org/10.1145/3405755.3406151

Dilmegani, C. (2026). Bias in AI: Examples and 6 Ways to Fix it in 2026. *AIMultiple*. [59] https://aimultiple.com/ai-bias

Ertel, W. (2016). Maschinelles Lernen und Data Mining. In *Grundkurs künstliche intelligenz. Computational Intelligence* (4th ed.). Springer Vieweg. [9] https://doi.org/10.1007/978-3-658-13549-2_8

European Union. (2024). Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending certain legislative acts. *Official Journal of the European Union*, L 1689. [10] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L_202401689

García-Micó, T. G., & Laukyte, M. (2023). Gender, health, and AI: How using AI to empower women could positively impact the sustainable development goals. In F. Mazzi & L. Floridi (Eds.), *The ethics of artificial intelligence for the sustainable development goals* (Philosophical Studies Series, Vol. 152, pp. 291–304). Springer. [10] https://doi.org/10.1007/978-3-031-21147-8_16

Glickman, M., & Sharot, T. (2025). How human–AI feedback loops alter human perceptual, emotional and social judgements. *Nature Human Behaviour*, 9, 345–359. https://doi.org/10.1038/s41562-024-02077-2

Guilbeault, D., Delecourt, S., & Desikan, B. S. (2025). Age and gender distortion in online media and large language models. *Nature*, 646(8087), 1129–1137. https://doi.org/10.1038/s41586-025-09581-z

Habler, F., Schwind, V., & Henze, N. (2019). Effects of smart virtual assistants' gender and language. In *Proceedings of Mensch und Computer 2019 (MuC '19)* (pp. 469–473). Association for Computing Machinery. [11] https://doi.org/10.1145/3340764.3344441

Hansen, M., Schwartmann, R., & Köhler, M. (2023, October 22). Was gegen die Voreingenommenheit der KI-Systeme hilft. *Frankfurter Allgemeine Zeitung*. [11] https://www.faz.net/aktuell/wirtschaft/unternehmen/was-gegen-die-voreingenommenheit-der-ki-systeme-hilft-19260960.html

Harari, Y. N. (2017). *Homo deus: A history of tomorrow*. Harvill Secker. [12]

Ho, J. Q. H., Hartanto, A., Koh, A., & Majeed, N. M. (2025). Gender biases within Artificial Intelligence and ChatGPT: Evidence, sources of biases and solutions. *Computers in Human Behavior: Artificial Humans*, 4, Article 100145. [41] https://doi.org/10.1016/j.chbah.2025.100145

Independent Group of Scientists appointed by the Secretary-General. (2023). *Global sustainable development report 2023: Times of crisis, times of change: Science for accelerating transformations to sustainable development*. United Nations. [12]

Inter-Parliamentary Union & UN Women. (2025). *Women in politics: 2025*. [12] https://www.unwomen.org/sites/default/files/2025-03/women-in-politics-2025-en.pdf

International Telecommunication Union. (2020). *United Nations activities on artificial intelligence (AI): 2020*. [13] https://www.itu.int/dms_pub/itu-s/opb/gen/S-GEN-UNACT-2020-1-PDF-E.pdf

International Telecommunication Union. (2024). *Facts and figures 2024 – The gender digital divide*. [13] https://www.itu.int/itu-d/reports/statistics/2024/11/10/ff24-the-gender-digital-divide/

Jabbour S, Fouhey D, Shepard S, et al. (2023) *Measuring the impact of AI in the diagnosis of hospitalized patients: A randomized clinical vignette survey study. JAMA.* 2023;330(23):2275-2284. doi:10.1001/jama.2023.22295

Josten, C., & Lordan, G. (2020). Robots at work: Automatable and non-automatable jobs. In U. K. Zimmermann (Ed.), *Handbook of labor, human resources, and population economics* (pp. 1–24). Springer. [13] https://doi.org/10.1007/978-3-319-57365-6_10-1

Karam, H., Cai, M., Palliyil, V., Phatthanachaisuksiri, L., Suhre, N., & Kaßens-Noor, E. (2025). Mapping AI's role in supporting or hindering the achievement of SDGs: A systematic review for AI in government. *Open Research Europe*, 5, Article 101. [14] https://doi.org/10.12688/openreseurope.19696.1

Kay, M., Matuszek, C., & Munson, S. (2015). Unequal representation and gender stereotypes in image search results for occupations. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)* (pp. 3819–3828). Association for Computing Machinery. [14] https://doi.org/10.1145/2702123.2702520

Khor, S., Haupt, E. C., Hahn, E. E., Lyons, L. J. L., Shankaran, V., & Bansal, A. (2023). Racial and ethnic bias in risk prediction models for colorectal cancer recurrence when race and ethnicity are omitted as predictors. *JAMA Network Open*, 6(6), e2318495. https://doi.org/10.1001/jamanetworkopen.2023.18495

Korytnikova, E. (2023). Artificial intelligence and women's health: Innovations, challenges, and ethical considerations. *Advances in Clinical and Medical Research*, 4(3), 1–6. [15] https://doi.org/10.52793/ACMR.2023.4(3)-59

Laffier, J., & Rehman, A. (2023). Deepfakes and harm to women. *Journal of Digital Life and Learning*, 3(1), 1–21. [15] https://doi.org/10.51357/jdll.v3i1.218

Lenzen, M. (2020). *Künstliche Intelligenz: Fakten, Chancen*. Springer. [16]

Leslie, D. (2019). *Understanding artificial intelligence ethics and safety: A guide for the responsible design and implementation of AI systems in the public sector*. Zenodo. [16] https://doi.org/10.5281/zenodo.3240529

Liu, A. T. H., Lee, M., & Bai, X. (2026). Writing with AI can reduce gender bias in hiring evaluations. In *Proceedings of the 2026 CHI Conference on Human Factors in*

*Computing Systems*. Association for Computing Machinery. [92] https://doi.org/10.1145/3772318.3791136

Lütz, F. (2023). Gender equality and artificial intelligence: SDG 5 and the role of the UN in fighting stereotypes, biases, and gender discrimination. In E. Fornalé & F. Cristani (Eds.), *Women's empowerment and its limits* (pp. 1–16). Palgrave Macmillan. [16] https://doi.org/10.1007/978-3-031-29332-0_9

Martin, K. D., Borah, A., & Palmatier, R. W. (2017). Data privacy: Effects on customer and firm performance. *Journal of Marketing*, 81(1), 36–58. [17] https://doi.org/10.1509/jm.15.0497

Mayoyo, N. (2023). The influence of social media use in the wake of deep fakes on Kenyan female university students' perceptions on sexism, their body image, and participation in politics. In K. Langmia (Ed.), *Black communication in the age of disinformation* (pp. 89–103). Palgrave Macmillan. [17] https://doi.org/10.1007/978-3-031-27696-5_5

Moore, P. V. (2019). OSH and the future of work: Benefits and risks of artificial intelligence tools in workplaces. In V. Duffy (Ed.), *Digital human modeling and applications in health, safety, ergonomics and risk management* (LNCS Vol. 11581, pp. 234–245). Springer. [18] https://doi.org/10.1007/978-3-030-22216-1_22

OECD. (2021). *Pay transparency tools to close the gender wage gap*. [19] https://www.oecd.org/en/publications/pay-transparency-tools-to-close-the-gender-wage-gap_eba5b91d-en.html

Omdena. (2025). *Preventing sexual harassment with artificial intelligence*. [19] https://www.omdena.com/projects/sexual-harassment/

Pal, S., Lazzaroni, R. M., & Mendoza, P. (2024). AI's missing link: The gender gap in the talent pool. *Interface*. [19] https://www.interface-eu.org/publications/ai-gender-gap

Palomares, I., Martínez-Cámara, E., Montes, R., García-Moral, P., Chiachio, M., Chiachio, J., & Herrera, F. (2021). A panoramic view and SWOT analysis of artificial intelligence for achieving the sustainable development goals by 2030: Progress and prospects. *Applied Intelligence*, 51, 6497–6527. [20] https://doi.org/10.1007/s10489-021-02264-y

Pasquale, F. (2015). *The black box society: The secret algorithms that control money and information*. Harvard University Press. [20]

Patón-Romero, J. D., Vinuesa, R., Jaccheri, L., & Baldassarre, M. T. (2022). State of gender equality in and by artificial intelligence. *IADIS International Journal on Computer Science and Information Systems*, 17(2), 31–48. [21]

Pigola, A., Da Costa, P. R., Carvalho, L. C., Silva, L. F. D., Kniess, C. T., & Maccari, E. A. (2021). Artificial intelligence-driven digital technologies to the implementation of the

sustainable development goals: A perspective from Brazil and Portugal. *Sustainability*, 13(24), Article 13669. [21] https://doi.org/10.3390/su132413669

Pokhriyal, R., Bahuguna, R., Memoria, M., & Kumar, R. (2023). ICT facilitated domestic violence in India: A quantitative study on experts' perception. In *2023 International conference on advancement in computation and computer technologies (InCACCT)* (pp. 134–138). IEEE. [22] https://doi.org/10.1109/InCACCT57535.2023.10141824

Raj, A., Dehingia, N., Singh, A., McDougal, L., & McAuley, J. (2020). Application of machine learning to understand child marriage in India. *SSM - Population Health*, 12, Article 100687. [22] https://doi.org/10.1016/j.ssmph.2020.100687

Rane, N. (2023). *Roles and challenges of ChatGPT and similar generative artificial intelligence for achieving the sustainable development goals (SDGs)*. SSRN 4603244. [23] https://doi.org/10.2139/ssrn.4603244

Rodriguez-Davila, N. (2025). Machine learned misogyny: Gender bias in AI. *DePaul Journal of Art, Technology & Intellectual Property Law*, 35, 71–123. [85]

Russell, S., & Norvig, P. (2020). *Artificial intelligence: A modern approach* (4th ed.). Pearson. [23]

Sharma, P. (2024). Artificial intelligence as a catalyst for gender equality: Empowering women across healthcare, education, and entrepreneurship. *Mind & Society*, 13(3), 34–41. [23] https://doi.org/10.56011/mind-mri-133-20245

Shrestha, B., Luvanda, A., & Razzaq, J. (2026, May). *Feminist and Global South perspectives on AI-supported learning environments*. Brookings Commentary. Brookings Institution. [70] https://www.brookings.edu/blog/education/2026/05/feminist-global-south-perspectives-ai/

Shumailov, I., Shumaylov, Z., Zhao, Y., Gal, Y., Papernot, N., Anderson, J. H., & McGillivray, R. J. (2024). AI models collapse when trained on recursively generated data. *Nature*, 631, 755–759. https://doi.org/10.1038/s41586-024-07566-y

Smith, G., & Rustagi, I. (2021). When good algorithms go sexist: Why and how to advance AI gender equity. *Stanford Social Innovation Review*. https://ssir.org/articles/entry/when_good_algorithms_go_sexist_why_and_how_to_advance_ai_gender_equity

Sperlic, S., Heller, C., Geyer, S., & Ttzlaff, F. (2024). Living longer, working longer: Analysing time trends in working life expectancy in Germany from a health perspective between 2002 and 2018. *European Journal of Ageing*, 19(1), 1263–1276. [24] https://doi.org/10.1007/s10433-022-00707-0

Sumathy, B., Shiva, P., Mugundhan, P., Rakesh, R., & Prasath, S. (2019). Virtual friendly device for women security. *Journal of Physics: Conference Series*, 1362, 12042. [24] https://doi.org/10.1088/1742-6596/1362/1/012042

UN Women. (2023). *Forecasting time spent in unpaid care and domestic work*. [26] https://data.unwomen.org/publications/forecasting-time-spent-unpaid-care-and-domestic-work

UN Women. (2025). *Women's rights in review: 30 years after Beijing*. [26] https://www.unwomen.org/sites/default/files/2025-03/womens-rights-in-review-30-years-after-beijing-en.pdf

UN Women. (2026, April). *AI is already rewriting reality for billions of people. It is getting women wrong*. UN Women Press Advisory. https://www.unwomen.org/en/news-and-events/stories/2026/04/ai-bias-rewriting-reality

UN-Women & DESA. (2024). *Gender data outlook 2024 world bank, statistical performance indicators SDG 5 tracker united nations*. [27]

UNCTAD. (2018). *United Nations Conference on Trade and Development: World investment report 2018*. [25] https://unctad.org/en/PublicationsLibrary/wir2018_en.pdf

United Nations. (2024). *The sustainable development goals report 2024*. [25] https://unstats.un.org/sdgs/report/2024

United Nations Department of Economic and Social Affairs. (2025). *World population prospects 2024*. [25] https://population.un.org/wpp

United Nations General Assembly. (2015). *Transforming our world: The 2030 agenda for sustainable development* (A/RES/70/1). [25] https://undocs.org/A/RES/70/1

van den Broek, E., Sergeeva, A. V., & Huysman, M. (2025a). Is there fairness in AI? *Journal of Management Studies*, Advance online publication. https://doi.org/10.1111/joms.13276

van den Broek, E., Sergeeva, A. V., & Huysman, M. (2025b). New research on AI and fairness in hiring. *Harvard Business Review*, 103(6). https://hbr.org/2025/11/new-research-on-ai-and-fairness-in-hiring

Vinuesa, R., Azizpour, H., Leite, I., Balaam, M., Dignum, V., Domisch, S., Felländer, A., Langhans, S. D., Tegmark, M., & Nerini, F. (2020). The role of artificial intelligence in achieving the sustainable development goals. *Nature Communications*, 11(1), Article 233. [27] https://doi.org/10.1038/s41467-019-14108-y

Vorvoreanu, M., Zhang, L., Huang, Y.-H., Hilderbrand, C., Steine-Hanson, Z., & Burnett, M. (2019). From gender biases to gender-inclusive design: An empirical investigation. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp.

1–14). Association for Computing Machinery. [28] https://doi.org/10.1145/3290605.3300283

Walsh, T. (2018). *It's alive: Wie künstliche Intelligenz unser Leben verändern wird*. Edition Körber. [28]

Wang, Z., Wu, Z., Zhang, J., Guan, X., Jain, N., Lu, S., Gupta, S., & Koshiyama, A. (2024). *Bias amplification: Large language models as increasingly biased media*. arXiv preprint arXiv:2410.15234. https://doi.org/10.48550/arXiv.2410.15234

West, M., Kraut, R., & Chew, H. E. (2019). *I'd blush if I could: Closing gender divides in digital skills through education*. UNESCO Policy Brief. https://unesdoc.unesco.org/ark:/48223/pf0000367416

World Economic Forum. (2024). *Global gender gap report 2024*. WEF. [29] https://www3.weforum.org/docs/WEF_GGGR_2024.pdf

World Wide Web Foundation. (2015). *Women's rights online: Translating access into empowerment*. [29] https://webfoundation.org/research/womens-rights-online-2015/

Zack, T. (2024). Assessing the potential of GPT-4 to perpetuate racial and gender biases in health care: A model evaluation study. *The Lancet Digital Health*, 6(1), e12–e22. https://doi.org/10.1016/S2589-7500(23)00225-X

Zweig, K. (2019). *Ein Algorithmus hat kein Taktgefühl*. Heyne. [29]