

AI Agency through the Lens of Data Feminism

Yu-Shien Yang (3210131)
OCAD University, Toronto

I. INTRODUCTION

In "A Cyborg Manifesto," Donna Haraway envisioned a post-gender world where the boundaries between human and machine blur, potentially liberating us from traditional gender constraints. Haraway optimistically proposed that cyborgs would transcend gender, opening a new era of fluid identities. (Haraway 1991)

However, the reality has diverged significantly from this utopian vision. As AI technology advances, we are not witnessing the de-gendering of AI humanoid robots or virtual assistants. Instead, gender biases are becoming increasingly pronounced in the digital world. This phenomenon reflects cultural and social norms that are being encoded into the technology itself. The design and deployment of AI technologies often mirror the biases of their creators and the societies in which they are developed.

This paper explores the complex interplay between AI agencies and gender, using the framework of data feminism to provide a critical lens on these issues. Data feminism emphasizes the importance of considering power dynamics in data science and technology, advocating for a more inclusive approach that recognizes and addresses inequalities. By applying this framework, the paper aims to uncover how AI technologies perpetuate gender biases and explore potential pathways for creating more equitable digital futures.

2. AI AGENCY AND GENDER

2.1 Technology and Gender in Cultural Context

The intersection of AI technology and gender roles presents a complex landscape where traditional values and rapid technological advancements create unique challenges and opportunities for representation. The portrayal of gender in media, such as science fiction anime and films, frequently depicts cyborg female characters with exaggerated physical features like large breasts and slender bodies (Sato 2004, 335-355). These portrayals reinforce stereotypical ideals of femininity and reflect social norms of gender and beauty.



Figure 1. The anime "Ghost in the Shell" was inspired by Donna Haraway's "A Cyborg Manifesto," yet in the animation, we still observe that the main character, Motoko Kusanagi, displays a very distinct gender characteristics such as huge boobs. 1995.

These cultural representations extend into the realm of AI, where technologies often mirror social biases rather than challenge them. For example, AI-generated images and beauty contests have highlighted how AI systems can perpetuate traditional standards of beauty, particularly the preference for fair-skinned, large-eyed women (Mouriquand 2024). This phenomenon shows the deep-seated cultural biases that are often encoded into AI systems, whether consciously or unconsciously.



Figure 2. Famous Japanese artist Hajime Sorayama in "Sex Matter", clearly depicts gender expectations and pronounced gender characteristics 2020



Figure 3. Miss AI, First AI Beauty Contest, 2024

2.2 Data Bias and AI Agency

AI systems, while designed to be neutral, frequently embody gender biases in societies. Take virtual AI assistants as an example, they are often designed with traditionally feminine characteristics, such as soothing female voices and subservient personalities. These design choices reinforce social expectations that associate women with caregiving and support roles (Madgavkar 2021).

Gender biases in training data can profoundly influence how AI systems function, the interactions they facilitate, and the narratives they perpetuate. Data bias arises when the datasets used to train these systems are unrepresentative, often resulting from factors such as the underrepresentation of certain groups, cultural stereotypes embedded in data, and historical inequities in data collection (Leavy 2018). In the context of gender, these biases

frequently manifest in ways that restrict the diversity of female representation in AI-generated content. This significantly limits the agency of female characters in virtual environments, leading to stereotypical portrayals where women are seen as overly emotional, nurturing, or submissive, misrepresenting the diversity of women's experiences and reinforcing outdated gender norms.

Moreover, biased data can constrain the narratives generated by AI systems, limiting the scope of actions and decisions available to female characters and perpetuating the perception that women are less capable or influential than their male counterparts (Wajcman 2007). This issue can be particularly pronounced in cultural contexts like East Asia, where social expectations of women are stringent, and traditional views on femininity, beauty, and social roles are deeply ingrained in the data.

3. WHAT IS DATA FEMINISM AND WHY IS IT IMPORTANT?

Data feminism, proposed by Catherine D'Ignazio and Lauren F. Klein, is a framework designed to uncover and challenge the power imbalances inherent in data science and technology. This approach emphasizes the inclusion of diverse perspectives, particularly those of marginalized groups, in the design and implementation of data-driven systems. The seven principles of data feminism provide a foundational lens for examining gender issues in AI. (D'Ignazio and Klein 2020)

3.1 Examine Power

The principle of examining power calls for a critical analysis of the power structures embedded in data systems. In the context of AI, this involves scrutinizing who creates these systems, whose interests they serve, and how they replicate existing social hierarchies. By examining the power dynamics within AI technologies, we can identify how traditional gender hierarchies are either challenged or reinforced in virtual environments. The authors also advocate for a shift from concepts like "fairness" and "accountability" to more transformative concepts like "equity" and "co-liberation," highlighting the need for data practices that address root causes of inequality rather than just surface-level issues (D'Ignazio and Klein 2020, 21-47).

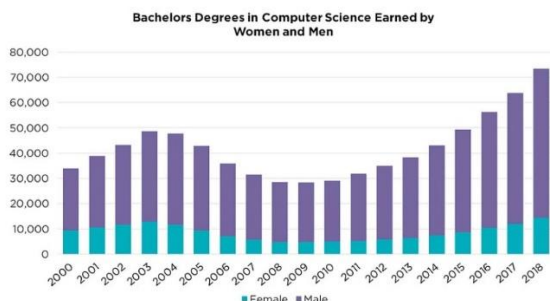


Figure 4. Computer Science is a Man Factory, (Code.org 2020)

3.2 Challenge Power

Building on the examination of power, this principle urges a proactive approach to challenge and disrupt unjust power dynamics. It emphasizes the importance of using data science to confront and alter existing power structures. In AI, this means advocating for technologies that do not merely replicate social biases but actively work against them. Challenging power in AI systems includes questioning the default assumptions and design choices that often

perpetuate gender stereotypes, particularly in the representation of virtual assistants and robots (D'Ignazio and Klein 2020, 49-72). In practice, we can achieve this goal by collecting counterdata and including more female and LGBTQ+ data scientists.

3.3 Elevate Emotion and Embodiment

Data feminism emphasizes the importance of acknowledging emotion and embodiment, countering the traditional view that data science should be objective and emotionless. In AI, this principle is particularly relevant to the design and functionality of gendered AI agents. It encourages designers to consider how these agents embody emotional and physical characteristics and how such embodiments can perpetuate or challenge gender norms. In practice, the authors introduce the concept of "data visceralization," a method that goes beyond visualization to engage multiple senses in experiencing data, helping people better understand and remember the stories behind the data (D'Ignazio and Klein 2020, 73-96).

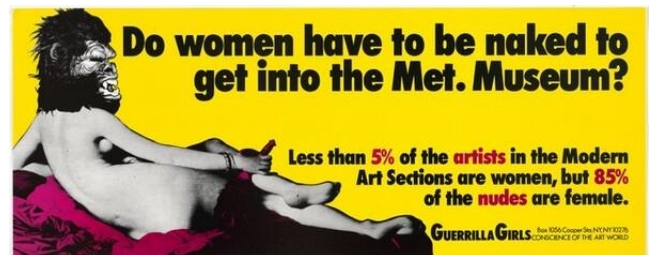


Figure 5. Do Women Have to Be Naked to Get into the Met. Museum? A data-driven infographic created by the Guerrilla Girls in 1989.

3.4 Rethink Binaries and Hierarchies

This principle invites us to question the assumptions behind classification systems, such as human/machine, male/female. The gender binary is socially constructed and often fails to reflect the true diversity of human experiences. In AI, these binaries are manifested in the clear-cut gender representations of virtual agents and robots. By rethinking these binaries, we can foster more diverse representations of gender, moving beyond simplistic and often stereotypical portrayals that better reflect the complexities of human identities and experiences. (D'Ignazio and Klein 2020, 98-123).

3.5 Embrace Pluralism

Embracing pluralism involves recognizing and valuing diverse perspectives and experiences. In the context of AI, it calls for the inclusion of varied representations of gender and identity, challenging the monolithic views often presented in technology. This principle supports the creation of AI systems that reflect a wider range of human experiences, thereby fostering inclusivity (D'Ignazio and Klein 2020, 125-148).

3.6 Consider Context

The authors argue that data are not neutral or objective; they are produced within unequal social relations and thus carry the biases and power dynamics of their context. The principle of considering context emphasizes the importance of understanding the specific cultural, social, and historical contexts in which data is produced and interpreted. In AI, context is crucial for understanding how gender biases are encoded and perpetuated in different cultural settings. This principle helps uncover the social norms that influence AI design and deployment, highlighting the need for

culturally sensitive approaches. The authors advocate for practices such as "data biographies" and "datasheets for datasets," which detail the provenance, collection methods, and limitations of data sets. These practices help users understand the data's context, ensuring more ethical and accurate analysis and communication (D'Ignazio and Klein 2020, 149-172).

3.7 Make Labor Visible

Finally, making labor visible involves acknowledging the often-invisible work that goes into creating and maintaining data systems. The authors argue that much of the work behind data products, such as data collection, cleaning, and analysis, as well as emotional and care labor, is often overlooked or undervalued.

The book highlights the example of "Anatomy of an AI System," a project that traces the human labor, data dependencies, and material resources involved in creating an Amazon Echo device. This project illustrates the vast and often hidden labor required at various stages, from mineral extraction to data processing. In the AI industry, this principle highlights the roles of individuals who design, program, and train AI systems. It also draws attention to the gendered nature of this labor, revealing how certain tasks are feminized and devalued (D'Ignazio and Klein 2020, 173-201).

The seven principles intersect with the concept of AI agents in several ways. It provides a critical perspective on gender issues in AI. Data feminism calls for a reimagining of data science and artificial intelligence beyond existing power structures, advocating for early resistance and reimagination before norms and regulations become fully established.

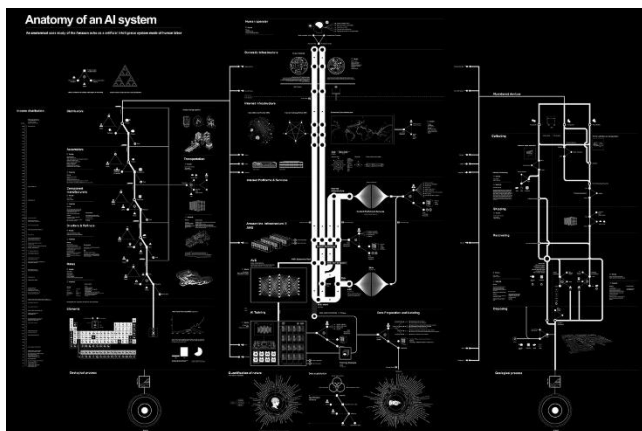


Figure 6. Crawford, Kate, and Vladan Joler, *Anatomy of an AI System*, 2018

4. EMPOWERMENT AND RESISTANCE

4.1 Reshaping Female Agency in AI

While the seven principles offer pathways to mitigate data bias, the challenges still remain. The complexity of cultural and social norms, especially in regions like East Asia, where gender biases are often deeply embedded and difficult to uproot. The deeply ingrained nature of cultural biases means that merely changing the design of AI systems is not sufficient. There must also be a broader social shift in the perception and valuation of diverse gender identities and roles. This involves not only redesigning AI systems to avoid reinforcing stereotypes but also actively developing AI characters with complex personalities and diverse body types, or creating virtual spaces that encourage exploration of non-traditional gender

roles. For example, AI developers can consciously choose to diversify the representation of virtual assistants, incorporating a range of voices, appearances, and characteristics. Additionally, incorporating gender-sensitivity training for those involved in AI development can help mitigate unconscious biases that might otherwise be encoded into these technologies.



Figure 7. Humanoid Robot in Japan, 2018

Furthermore, technology provides powerful tools to enhance female agency. For example, virtual reality and augmented reality technologies can create spaces where traditional gender roles can be questioned and reimagined. AI-driven storytelling platforms offer the opportunity to craft narratives that challenge stereotypes and showcase diverse representations of women.

Artistic creation and technological innovation also play a crucial role in challenging gender stereotypes. When used intentionally, AI art generators can produce images that subvert traditional beauty standards. Additionally, feminist AI collectives and artists are emerging, focusing on developing AI systems that prioritize gender equality and diverse representation.



Figure 8. Mariko Mori, Tea Ceremony, 2018

4.2 The Role of "Glitch Feminism"

While the current landscape of AI may seem to contradict Donna Haraway's optimistic vision of a post-gender world, there still is substantial potential for transformative change. As Legacy Russell suggests, the glitches and imperfections in our virtual world can serve as opportunities for resistance and reimagination. These disruptions provide critical moments where existing norms can be

questioned. It provides a framework for understanding and utilizing technological errors or "glitches" as moments of resistance against programmed norms. Glitch feminism celebrates these disruptions as opportunities to challenge and subvert traditional gender roles (Russell 2020).

In the context of AI, glitches can manifest as unexpected behaviors or outputs that do not conform to gendered expectations. These moments of unpredictability offer a space for reimagining female agency in AI systems. By embracing non-traditional representations and behaviors, glitch feminism advocates for a broader, more inclusive understanding of gender in technology. This approach also reminds us about Hito Steyerl's "In Defense of the Poor Image", which discusses how poor images often bypass traditional channels of distribution, allowing for broader access and democratization of visual culture (Steyerl 2009).



Figure 9. Legacy Russel, *Glitch Feminism*, 2020

5. CONCLUSION

Data feminism and glitch feminism provide us a vital framework for understanding and addressing issues of female agency in AI. By applying its principles, we can uncover hidden biases, challenge entrenched power structures, and strive for more equitable representations in virtual spaces. This framework not only highlights the limitations of current AI systems but also offers a roadmap for creating technologies that reflect a broader spectrum of human experiences.

The study of gender issues in AI technologies is crucial, as AI increasingly shapes our social and cultural landscapes. The portrayal and treatment of gender in AI technologies are not merely local concerns, they resonate globally, influencing standards and expectations across different societies.

Looking to the future, continued research is essential in several key areas. Developing culturally sensitive AI ethics frameworks is paramount to ensure that AI technologies respect and reflect diverse values and norms. There is also a need to design AI systems that actively promote gender equality, incorporating inclusive design practices, bias detection algorithms, and educational initiatives. AI-driven virtual spaces offer significant potential as platforms for gender education and empowerment, providing new ways to challenge and transform social norms.

The emergence of cyborg women is an emergency for men in the 21 centuries. However, the future is unmanned. Gender roles and identities can be fluid, diverse, and liberated from conventional constraints. The study of gender and AI not only reveals the

complexities and challenges of this intersection but also offers a glimpse into a more inclusive future. The path forward requires vigilance, creativity, and a commitment to equity, ensuring that AI technologies serve all of humanity, not just a privileged few. By continuing to explore and address these issues, we can work towards a world where AI systems enhance rather than limit human potential, embracing the full diversity of gender identities and roles.

6. REFERENCES

- [1] Butler, Judith. *Gender Trouble: Feminism and the Subversion of Identity*. New York: Routledge, 1990.
- [2] Code.org. "Women Computer Science Graduates Finally Surpass Record Set 17 Years Ago, but Percentages Lag Behind." Medium, May 11, 2020. <https://codeorg.medium.com/women-computer-science-graduates-finally-surpass-record-set-17-years-ago-20a79a76275>.
- [3] Crawford, Kate, and Vladan Joler. *Anatomy of an AI System: The Amazon Echo as an Anatomical Map of Human Labor, Data, and Planetary Resources*. 2018. <https://anatomyof.ai/>.
- [4] D'Ignazio, Catherine, and Lauren F. Klein. *Data Feminism*. Cambridge, MA: The MIT Press, 2020.
- [5] Haraway, Donna J. "A Cyborg Manifesto: Science, Technology, and Socialist-Feminism in the Late 20th Century." In *Simians, Cyborgs and Women: The Reinvention of Nature*, 149-181. New York: Routledge, 1991.
- [6] Laboria Cuboniks. *The Xenofeminist Manifesto: A Politics for Alienation*. Illustrated. London: Verso, 2018.
- [7] Leavy, Susan. "Gender Bias in Artificial Intelligence: The Need for Diversity and Gender Theory in Machine Learning." In 2018 IEEE/ACM 1st International Workshop on Gender Equality in Software Engineering (GE), 14-16. IEEE, 2018.
- [8] Madgavkar, Anu. "A Conversation on Artificial Intelligence and Gender Bias." Podcast, April 7, 2021. McKinsey & Company. <https://www.mckinsey.com/featured-insights/asia-pacific/a-conversation-on-artificial-intelligence-and-gender-bias>.
- [9] Mouriquand, David. 2024. "Miss AI: World's First Beauty Contest with Computer Generated Women." Euronews, April 25, 2024. <https://www.euronews.com/culture/2024/04/25/miss-ai-worlds-first-beauty-contest-with-computer-generated-women>.
- [10] Russell, Legacy. *Glitch Feminism: A Manifesto*. London: Verso Books, 2020.
- [11] Saito, Kumiko. "Magic, 'Shōjo', and Metamorphosis: Magical Girl Anime and the Challenges of Changing Gender Identities in Japanese Society." *The Journal of Asian Studies* 73, no. 1 (February 2014): 143-164. Association for Asian Studies. <https://www.jstor.org/stable/43553398>.
- [12] Sato, Kumiko. "How Information Technology Has (Not) Changed Feminism and Japanism: Cyberpunk in the Japanese Context." *Comparative Literature Studies* 41, no. 3 (2004): 335-355. Penn State University Press. <https://www.jstor.org/stable/40247417>.
- [13] Steyerl, Hito. "In Defense of the Poor Image." *e-flux Journal*, no. 10 (November 2009). <https://www.e-flux.com/journal/10/61362/in-defense-of-the-poor-image/>.
- [14] Suchman, Lucy. "Demystifying the Intelligent Machine." In *Cyborg Futures*, November 2019. https://doi.org/10.1007/978-3-030-21836-2_3.
- [15] Wallis, Jonathan. "The Paradox of Mariko Mori's Women in Post-Bubble Japan: Office Ladies, Schoolgirls, and Video-Vixens." *Woman's Art Journal* 29, no. 1 (Spring-Summer 2008): 3-12. Old City Publishing, Inc. <https://www.jstor.org/stable/20358141>.
- [16] Wajcman, Judy. "From Women and Technology to Gendered Technoscience." *Information, Communication & Society* 10, no. 3 (2007): 287-29.