Feminist AI: Critical Perspectives on Algorithms, Data, and Intelligent Machines

by Jude Browne, Stephen Cave, Eleanor Drage and Kerry McInerney (eds.) (2023) New York, Oxford University Press, 432 pp.

Ericka Johnson
Linköping University

feminist AI joins the growing number of books that examine the complexities and intricacies of AI's socio-technical entanglements, though with an avowedly feminist approach. Many of the volume's contributors have recently published their own, book length, manuscripts (D'Ignazio and Klein 2019; Costanza-Chock 2020), but the strength of *feminist AI* is in the bouquet of voices, theoretical perspectives, disciplinary approaches and methods that it brings together between two covers. It is a true *tour de force* demonstrating the analytical usefulness of feminist theory.

The chapters in *feminist AI* address a wide range of concerns – everything from asking if AI *can* be feminist to intersectional analyses of how the gendered effects of AI engage with other forms of power and control (such as ethnicity, race, class, disability, sexuality, and age). There is work on AI and labor (Chapter 11 *Of Techno-Ethics and Techno-Affects* by Sareeta Amrute), imaginaries of AI's potentially queer technological future (Chapter 9 *Feminist Technofutures: Contesting the Ethics and Politics of Sex Robots and AI* by Neda Atanasoski), examples of how alternative epistemologies can be engaged (Chapter 17 *Automating Autism* by Os Keyes) and what we mean by human (Chapter 6 *No Humans in the Loop: Killer Robots, Race and AI* by Lauren Wilcox).

The book presents intersectional feminism as a political project challenging social injustices, with example after example of how feminism is particularly well positioned to approach the risks and benefits of AI. But the collection of chapters in this anthology also displays the usefulness of a multidisciplinary approach to AI. The many authors include academics working in the humanities, social sciences and the computer sciences, but also contributors who are primarily working within community activism, the tech industry and the arts. In contrast to common authorship practices within the humanities and social sciences, a good number of the chapters are co-authored, reflecting interdisciplinary research constellations. Yet, all of the work in this volume employs lessons from critical theory, in particular a sensitivity to the power dynamics AI is entangled within, which creates a clear sense of cohesion for the reader.

TECNOSCIENZA. Italian Journal of Science & Technology Studies
16(1) pp. 135-138 ISSN 2038-3460
DOI: 10.6092/issn.2038-3460/22438

While AI is often framed as very new and cutting edge, this book refuses to fall into a discourse wrapped in an ahistorical present. Rather, there are contributions from foundational thinkers within the field of gender & technology (for example, Chapter 1 *Technosymbiosis: Figuring (Out) Our Relations to AI* by N. Katherine Hayles and Chapter 4 *Feminism Confronts AI: The Gender Relations of Digitalisation* by Judy Wajcman and Erin Young) that build on early lessons about the technology & gender relation and apply them *con gusto* to AI (see also Chapter 3 *AI in a Different Voice: Rethinking Computers, Learning, and Gender Difference at MIT in the 1980s* by Apolline Taillandier which explores an early attempt to create feminist coding and Chapter 15 *The Cruel Optimism of Technological Dreams: Thinking AI through Lauren Berlant* by Caroline Bassett that engages Berlant's concept of Cruel Optimism). These chapters, which use traditional theoretical approaches are joined by other chapters from scholars working in and with decolonial contexts and epistemologies (see for example Chapter 2 *Making Kin with the Machines* by Jason Edward Lewis, Noelani Arista, Archer Pechawis, and Suzanne Kite and Chapter 20 *Afrofeminist Data Futures* by Neema Iyer, Garnett Achieng and Chenai Chair).

Feminist technoscience has a long history of engaging with cutting edge technology through a feminist lens sensitive to power, and many of the early lessons we learned in the field are relevant today, too. For example, early conversations about the posthuman and our relations with technology that N. Katherine Hayles developed are also returned to and engaged in the first chapter, *Technosymbiosis: Figuring (Out) Our Relations to AI*. Hayles engages very high-volume questions, reminding the reader about the value of early feminist technoscience staples like the concept of the "informatics of domination" in Haraway's Cyborg Manifesto (1985/1991). Hayles also challenges the reader to think about the unique and evolving capacities of AI and theorize about power in ways that capture the technological specificities of algorithmic processes in AI while still "sticking with the trouble" of situatedness and power. Calling this an ontological approach, Hayles urges us to use feminist strategies like attending to metaphors and considering "othered" beings, but also to do the political work of building alliances, collaborations and affinities with people (often men) already working with AI so that we can make positive interventions to address the goals and assumptions of, for example, predictive algorithms (p. 12). Judy Wajcman and Erin Young also bring some of the early lessons from feminist work with technology to AI (Chapter 4, *Feminism Confronts AI: The Gender Relations of Digitalisation*). We are reminded of the importance of thinking about who (and whose bodies, subjectivities, situated knowledges) is engaged in making the technology and how that impacts the supposedly "neutral" technologies they make (p. 58). Here Wajcman and Young also push back at the attempts from within computer science to mitigate bias, problematizing the idea that there could be an "objective" measure of fairness or even of data, and again referencing early work critiquing the "view from nowhere" that these approaches embrace.

Finally, feminist critiques of well-known AI problems abound in the book – rightly so but with often quite nuanced analyses. The reader is treated to in-depth discussions about predictive policing and gender-based violence (Chapter 7, *Coding "Carnal Knowledge" into Carceral Systems: A Feminist Abolitionist Approach to Predictive Policing* by Kerry Mackereth), categorization practices, race and capitalism (Chapter 8, *Techno Racial Capitalism: A Decolonial Black Feminist Marxist Perspective* by Lelia Marie Hampton), digital assistants,

reproductive labor and care (Chapter 10, *From ELIZA to Alexa: Automated Care Labour and the Otherwise of Radical Care* by Jennifer Rhee), and the reappearance of physiognomy – defining types of humans from appearance – in AI technologies (Chapter 13 *Physiognomy's New Clothes* by Blaise Aguera y Arcas, Margaret Mitchell and Alexander Todorov). Chapters 13 and 14 (*Signs Taken for Wonders: AI, Art and the Matter of Race* by Michele Elam), in particular, provide a very good bridge between historical practices of categorization and racialization, academic work that has problematized these practices outside the field of AI, and the sudden impetus to embrace obviously problematic categorization practices again within AI because of the affordances of image recognition and ML.

A particular strength of the book is that the various chapters engage analyses of power in the multiple ways that feminism does, while not shying away from using the word feminism. *feminist AI* is a collection which shows how critical feminist theory can be productive for empirical studies of STS that could be considered distinctly feminist in that they are concerned with power and othering, but which are not necessarily labeled "feminist". The chapters in *feminist AI* bring feminism into conversation with relevant STS literatures and debates about how a critique of power is necessary to examine the social entanglements of AI. One hopes that the use of the word feminist in the title will not scare off colleagues in STS who should read these studies but don't consider feminist theory and intersectionality relevant to their work. Kudos to Tecnoscienza for reviewing this book, because there are many contributions in it which will be useful for STS studies, especially those working in areas that overlap with critical data studies, the medical humanities, and design.

The book's primary weakness was that only Chapters 5 (*Shuri in the Sea of Dudes: The Cultural Construction of the AI Engineer* by Stephen Cave, Kanta Dihal, Eleanor Drage and Kerry Mackereth), 12 (*The False Binary of Reason and Emotion in Data Visualisation* by Catherine D'Ignazio and Lauren Klein), and 18 (*Digital Ageism, Algorithmic Bias and Feminist Critical Theory* by Rune Nyrup, Charlene Chu and Elena Falco) were initially open access. Luckily, this has changed and now it appears as if all chapters are available on the Oxford University Press site[1]. This is great because especially Chapters 10, 13, 14 and 19 (*AI & Structural Injustice: A Feminist Perspective* by Jude Browne) – will be useful in the classroom, and open access literature is often a necessity for many less privileged institutions and students. This availability will also make the exciting conversations in the book much more accessible, especially to a global audience.

Throughout *feminist AI*, we as readers are reminded of the relevance of much work within feminist theory which – broadly – points to how we and the world are created and performed through the very processes of "collecting" and "categorizing" data. The details of how that collecting and categorizing our ways of being in the world happen, and how they become the material used for learning and training AI, matter. They matter because those practices of seeing us and our lives in very particular ways create very particular data representations. Many of the chapters in the book discuss the insight that, through methods of collecting and categorizing, we become legible to the technological structures and artifacts that afford, disafford and dys-afford ways of being (see in particular Chapter 21 *Design Practices: Nothing About Us Without Us* by Sasha Costanza-Chock; see also Costanza-Chock 2020).

This insight is also extremely relevant when thinking about how AI systems use data to learn the world in iterative processes. It inspires one to revisit work on legibility and intelligibility

(Chapter 16, *AI that Matters: A Feminist Approach to the Study of Intelligent Machines* by Eleanor Drage and Federica Frabetti in particular). The large discussions about AI in the book are all nuanced with a feminist approach that sensitizes the analysis to embedded structures of power. This is an approach which reminds us that it is both important to analyze the technology, but equally important to pay attention to the details of its entanglements and the situated material and social structures that mediate or obfuscate AI's impacts (c.f. Chapters 18 and 19). Reading the book is both aggravating and inspiring – one is reminded of the current and potentially looming issues AI will bring to our practices – in the workplace, the market, education, our interactions with technological tools we use daily, etc., but also inspired to, like the many authors of this book, continue balancing on the narrow edge of analytical activism and engaging, collaborating and building alliances. One can see this collection as an intervention, an intervention I hope will be part of an ongoing conversation about how AI can be implicated in working for social justice within STS.

Notes

[1] https://academic.oup.com/book/55103.

References

Costanza-Chock, Sasha (2020) *Design Justice*, Cambridge (MA) and London (UK), MIT Press.

D'Ignazio, Catherine and Klein, Lauren (2019) *Data Feminism*, Cambridge (MA) and London (UK), MIT Press.

Haraway, Donna J. (ed.) (1991) (orig. pub. 1985) *A Cyborg Manifesto: Science, Technology, and Socialist-Feminism in the Late Twentieth Century*, in *Simians, Cyborgs and Women: The Reinvention of Nature*, New York, Routledge.