書評

Kate Crawford
Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence
Yale University Press, April 2021, 327 頁

Amelia Katirai[*]

Artificial intelligence (AI) is increasingly changing our lives, our societies and our planet. We may be unaware of these applications—or intrusions—of AI in our daily lives and their often far-stretching implications. A growing body of literature seeks to shed light on these implications through critical analyses of advances in the field of AI and calls for more cautious approaches to its development and implementation. These works attempt to respond to the pressure of technological determinism, which causes technological development to appear to be an onrushing stream which we are powerless to stem, and to a lack of appropriate regulatory brakes, which heighten the fraught ethical nature of AI.

Kate Crawford, senior principal research at Microsoft Research and co-founder of the AI Now Institute at New York University, makes a meaningful and unique contribution to this body of work with her newly published text, *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. As the title suggests, she has compiled an atlas, mapping out the varied aspects of the landscape around AI and providing a guide both to those familiar with the terrain and those new to it. Ultimately, she has aimed to produce a "theory of AI that accounts for the states and corporations that drive and dominate it, the extractive mining that leaves an imprint on the planet, the mass capture of data, and the profoundly unequal and increasingly exploitative labor practices that sustain it" (11). Her maps are often topographical (10), focusing on the hierarchies of power embedded in AI and often obscured.

This is one of Crawford's key ideas: when we consider AI, we must consider the relationships of power behind it; even more than ethics, *this* is the lens we must

[*] Graduate School of Human Sciences, Kyosei Studies, Osaka University
a.katirai@hus.osaka-u.ac.jp

use to examine AI. For this reason, she looks not only at the technical issues of AI which have been the focus of recent attention—such as biased datasets used to train AI and the faulty assumptions built into algorithms—but also at the broader issues around this, such as the role of commercial interests in driving development, and the planetary costs of new technologies. In this sense, Crawford is taking a *kyosei* approach to AI (though she does not describe it as such): she looks at the sometimes symbiotic but more often exploitative relationships between not only individuals and the nonhuman technologies designed and used, but the impact of this on relationships between humans, and on the nonhuman entities we share our living planet with. To do this, she cuts through the deterministic hype around AI and instead focuses on the implications of AI for the most vulnerable. She takes on topics ranging from the planetary degradation that results from the development of technologies such as AI, to the ethically, financially, and physically precarious circumstances of human beings making AI a reality. Behind this, she charts the role of political and military interests, which intersect with commercial interests to determine the course of AI development.

With a field as complex and multifaceted as AI, Crawford's aim to map these issues into a single, accessible text is ambitious, but she is successful. Just as a map cannot include every intricacy, Crawford too is not "claiming universality" (11). Still, she provides broad outlines of the issues that not only general readers but also academic ones should be tuned to. She does not aim to "complete a global atlas" as the "very idea invokes capture and colonial control" (13) but instead seeks to map out the issues and examples that shaped her thinking about AI. Her use of clear and relatable examples to illustrate particular issues helps equip the reader with a set of tools to better conceptualize issues in other domains. She takes the reader along, as though on field trips, to sites relevant to each of these issues. Comprehensively referenced, her text also functions as a guide to other literature on AI.

Crawford begins with the example of Clever Hans, a horse in the early 1900s who was believed to have the capacity to solve math problems, but who in fact was attuned to subtle clues from those asking the questions. Crawford draws this story—often used as a cautionary tale in psychology—into the area of AI, showing that we cannot always know what a particular AI model has learned and how it is producing

its outcomes. This is set up as the first of many instances in which AI is not what it appears to be and reflects the idea that "AI is neither artificial nor intelligent" (8), a striking concept, although not novel to the field. But what is AI? Crawford highlights the "shifting and plastic" (19) nature of the concept but uses it here "to talk about the massive industrial formation that includes, politics, labor, culture, and capital" (9) while using "machine learning" to refer to its technical aspects.

With this groundwork laid, Crawford then moves into one of her most striking chapters, "Earth." Given the urgent concerns over the rapidly deteriorating health of our planet, Crawford's focus here on the environmental issues around AI is welcome. She sheds light on how we are increasingly reliant on technologies such as the "cloud," which appear to be abstract entities, but in fact, are rooted in material extraction. As she writes,

> Each object in the extended network of an AI system, from network routers to batteries to data centers, is built using elements that required billions of years to form inside the earth. From the perspective of deep time, we are extracting Earth's geological history to serve a split second of contemporary technological time, building devices like the Amazon Echo and the iPhone that are often designed to last for only a few years. (30)

As Crawford argues, the data economy itself is "premised on maintaining environmental ignorance" (42). This is visible, for example, in the massive emissions produced through the development of AI, which often go unaddressed in discussions of the ethics of AI.

Crawford links extraction with exploitation and the issues faced by the labor force that makes these technologies possible in "Labor." She focuses in particular on workers who are constrained by the exigencies of an AI-powered economy, looking at "how humans are increasingly treated like robots and what this means for the role of labor" (57). The reader accompanies Crawford on a visit to an Amazon warehouse, drawing attention to the inhumane pressures placed on workers, who must adapt to the excessively exacting timelines of the digital economy. She focuses, too, on crowd workers—the workers behind the curtain who allow AI to masquerade as fully automated when in fact it is far more dependent on human labor than it appears. It is

these workers who make it possible for AI to function while being neither artificial nor intelligent, as Crawford argues.

Where does the data that makes AI possible come from, then? This is Crawford's next focus, as she examines the complex politics behind the datasets used in developing AI. Data mining is the "new oil" (107), as data is increasingly captured on every aspect of our lives. At the same time, there has been a "rhetorical move that shifted the notion of data away from something personal, intimate, or subject to individual ownership and control toward something more inert and nonhuman" (113). These shifts mean that those working on developing AI are often too distant from those whose data they are reliant on. This gap facilitates the use of data that should be seen as unethical.

Data issues tie into Crawford's next chapter on "Classification." Recent years have seen a growing recognition of the potential for AI to produce results that are discriminatory and which reflect the underlying inequalities of our society. There is often a narrow emphasis on technological solutions through, for example, improving the quality of the data used without fundamentally rethinking the assumptions underpinning the technology. As Crawford suggest, what we must consider is, "what unspoken social and political theories underlie and are supported by these classifications of the world?" (127).

Crawford expands on this through her focus on "Affect." She takes emotion detection technologies as a case study and explores the faulty social and psychological theories that underlie them. This industry, worth over 17 billion dollars (151), has applications in settings ranging from security to employment to healthcare. However, it is based on incorrect fundamental assumptions about the universality of human affective displays. These have been shored up with training databases of images of affective displays that do not reflect spontaneous human emotion. Crawford compares emotion detection to the discredited field of phrenology, with its "spurious claims allowed to stand in support of power" (177).

Zooming out again, Crawford takes a broader approach in her next chapter but continues her analysis of the political interests behind AI. She takes us to its very roots, looking at how military interests have played a role in the development of AI. A focus here is on defence projects such as the JEDI (Joint Enterprise Defense

Infrastructure) contract with the Pentagon. This was originally designed as Project Maven at Google, but the project was pushed out due to objections from employees, and instead adopted at Microsoft, Crawford's own employer. She extends this discussion into the role of AI in surveillance. She critiques how these uses for surveillance often target and further disadvantage the most vulnerable, including refugees and those caught in the criminal justice system.

The maps of Crawford's atlas are brought together in her final chapter, "Power," as she makes a case for the importance of looking at AI holistically and through a lifecycle approach. Here, she puts forward her key argument, that:

> To understand what is at stake, we must focus less on ethics and more on power. … Instead of glorifying company founders, venture capitalists, and technical visionaries, we should begin with the lived experiences of those who are disempowered, discriminated against, and harmed by AI systems (225).

She calls for a "renewed politics of refusal" (226) based on principles of justice. This refusal pushes back against technological determinism and leads to a fundamental questioning of the utility and necessity of developing AI. She asks the reader to reject the idea that these tools of power "are also fit to transform schools, hospitals, cities and ecologies, as though they were value neutral calculators that can be applied everywhere" (227). She ends with a coda, "Space," focusing on the plans for journeys into space by billionaires like Jeff Bezos, funded by exorbitant wealth gained at the expense of others.

Throughout her text, Crawford successfully synthesizes the literature on the social, ethical, and environmental implications of AI. She covers a broad range of emergent issues, while also shedding light on topics that have previously been underexamined. Notable here is her discussion of the environmental impact of AI. This will be of interest to readers with a *kyosei*-inspired perspective. Her text is accessible and inspires the reader to journey deeper and explore further the topics covered—features of a good atlas. The writing is clear and succinct and is highly readable not only for those with specialist interests in the area but also for readers from the broader public.

However, one area of concern is Crawford's positionality. Over the last year, we have increasingly seen experts on ethics at major tech companies finding themselves in precarious positions. Crawford's critique of AI is robust, but the reader may be left wondering whether Crawford's positionality as a researcher bankrolled by Microsoft has meant that some points are left underdeveloped. This is visible, for example, in Crawford's discussion of defence contracts, where Crawford discusses the outcry over Project Maven at Google at length, yet only briefly addresses the contract granted to Microsoft, without critically evaluating its implications (191). A more direct acknowledgement of her positionality would have been beneficial.

Furthermore, Crawford's positionality is also reflected in the broadly Northern (and particularly North American) focus of the book. Though Crawford's early disclaimer clarifies that this atlas will not be "global" (13), this is perhaps where the book does fall short of what the reader may expect from an atlas. The promise and peril of AI are likely to look different depending on where one finds oneself globally, and this is something that Crawford could have teased out in greater detail.

Nonetheless, Kate Crawford's *Atlas of AI* makes a meaningful contribution to the field. Though there is overlap with other recent texts on AI, Crawford has neatly synthesized recent debates, adding in a focus on the environment. Her text is perhaps most valuable in that it is comprehensive, drawing together perspectives on technology, ecology, and society which a *kyosei* perspective shows us to be inextricably linked but which are often viewed as separate domains of concern. It is ambitious to argue that any true vision for a future with AI must factor in all aspects covered in Crawford's atlas. Yet, as we are pressed on every side with evidence of our failures to pursue just, sustainable, and wise development, it must be inconceivable that we would settle for any less.