BOOK REVIEW

Ethics for People Who Work in Tech, M. Steen (2022) 216pp., €128 hardback, Chapman & Hall/ CRC Press, New York, ISBN 9780367543303

The first thing you notice about this book is the pile of white Lego bricks on the cover. Is this what is to come unless we change? A singularly white outlook? A wasteland of plastic? The ideal state? An ivory tower? A book that promises to bring ethics down from the ivory tower – in only 216 pages. This is quite a promise for a tiny tome, and almost certainly too large to deliver. Still, approaching the book with curiosity, I hoped it would provide sorely needed practical guidance. The last four years have seen a proliferation of data ethics and AI strategies, many of which have had little effect.

Recently, we have seen that various companies collecting data have deliberately propagated addiction in teenagers (Rosenblatt, 2023); tracked individuals publicly without their consent (Leins, 2023); enabled prosecutions for abortions (Tangalakis-Lippert, 2023), provided chatbots to teenagers at risk of self-harm without any disclosures (Xiang, 2023) – and these were just this week. And yet, Microsoft has done away with its ethics team.

> The conflict underscores an ongoing tension for tech giants that build divisions dedicated to making their products more socially responsible. At their best, they help product teams anticipate potential misuses of technology and fix any problems before they ship. But they also have the job of saying 'no' or 'slow down' inside organizations that often don't want to hear it – or spelling out risks that could lead to legal headaches for the company if surfaced [sic] in legal discovery. And the resulting friction sometimes boils over into public view. (Schiffer and Newton, 2023)

So why should anyone even care? And how can a little book make a difference to companies that want to race ahead and not document risks to avoid potential litigation?

The first thing I liked about this book was its framing of terminology. For those working across fields, it is essential to define what is meant as many terms (including 'ground truth', 'fairness' and 'explainability') have very different meanings in different fields. I also liked how the author situated and introduced himself in the discourse – our professions, locations, genders and all other aspects of ourselves shape the way we view the world and our ethics. Acknowledging the particular lenses we have is a powerful way to introduce not only oneself, but the significance of diversity in thinking about different technologies. I particularly loved the Introduction's desire to help people to flourish. This will mean different things for different people – and that is OK. I also appreciated the author not referring to responsible systems or data – it is people who are responsible, people who are ethical and people who create the tools needed to ensure that these values are embedded.

The entry point to this text is delightful, easy, self-aware and textured. It draws the reader in and demands curiosity, not positioning nor answers. It is wonderful in the simplicity with which it describes what it is not. Its analogy to learning scuba diving appealed directly to me as a former diver, but the framing of asking oneself where one is situated, what one brings to the table and requesting flexibility in vision and considerations is a beautiful beginning and radically different from the approach of those looking to solve problems with yet more data and systems.

Despite my objections to Bruno Latour, it was inevitable to use him to frame the discussion, and Steen makes him as accessible as he can be made. Having worked with data scientists for some years now, I can say that even the conversation around technology having embedded

values is often difficult, and one which is not well or consistently framed in educational institutions. My favourite text on this on this matter is *The Whale and the Reactor: A Search for Limits in an Age of High Technology* by Langdon Winner (2020), Steen is remiss in not citing Winner earlier in his book, but this could be my personal prejudices speaking. Winner's work frames these values beautifully and helps technologists (and others) understand the power structures around the building of systems, both hardware and software, and their legacy for politics and people. That said, Winner is included in a later chapter, but I still think a description of his work alongside that of Latour would have been extremely helpful for readers struggling to understand the power and politics of systems more broadly. All too often, systems are looked at individually, especially those that are publicly interfacing (such as ChatGPT). The interconnected systems we do not see can have far greater impact on us and our lives. These are virtually impossible to interrogate or challenge and ever-widening power gaps are created between those who understand these tools and those who do not.

The chapter on smart home devices is fantastic, moving the conversation from the public to the private, a space that is only beginning to be explored in research (e.g., Melbourne Social Equity Institute, n.d.). I loved the ironic examples given in this chapter, but I would have liked a few more. Ring cameras and remotely controlled devices – how is data from these shared? Beyond privacy, who controls them and who understands how they can affect people? The examples given are humorous, but real-world applications are increasingly being suggested for our homes. What does this surveillance do to us? There is a large body of literature around the Panopticon and how it changed behaviour. What makes us human if nothing is private? 'Radical evil', as Hannah Arendt (1951) described it, is 'making human beings as human beings superfluous'. In the past, humans have helped us to recognize our own emotions and we have learned from others. If machines take on this function, what is it that makes us human? If everything is documented, does this change how we engage in the world (Tranter and Bikundo, 2018)?

It was great to see reference to Indigenous peoples. I would have loved to have had more references to work by Indigenous people in this field. I also enjoyed the problems, explored in Chapter 14, though they might have been better positioned at the end of the book, perhaps with notes for teaching. And Chapter 15 would have made more sense up-front as comparative ethics, rather than tucked away at the back. Still, virtue ethics is often overlooked in these conversations and needed to be considered.

Finally, applying these concepts is where the power of a book like this lies. Rather than providing checklists, ideas for challenging projects in healthy ways are required. Talk of how to embed these requirements in policy and company goals, with measurable outputs and long-term alignment with the vision for the company, is missing for me. So, too, is an outline, including the significance of culture and the protection of whistleblowers, both of which have a significant impact and are required to support even the best-intentioned organizations with policies, procedures and frameworks for data ethics.

Further, many organizations will review these ethical considerations through risk assessment. However, the normative nature of risk frameworks means that certain impacts or harms might be overlooked (Kaminski, 2023) Principles, culture and people are as critical in the governance of AI systems as the law, frameworks and risk management tools themselves. This includes the ability to sound the alarm when things go wrong[1] and not to be silenced (Byrne, 2021). It also requires the ability to challenge the thinking of systems as magical and that of humans as likely to fail.

Steen's exploration of human rights is fantastic. New standards are all pegged to human rights. Many countries (Australia excluded) have human rights regulatory frameworks that require compliance. However, how these frameworks are interpreted in the light of rapidly changing technologies remains uncertain. Research is required to complement mounting case law.

---

[1] See the work of Nancy Leveson, listed at http://sunnyday.mit.edu/book.html, accessed October 2023.

More generally (and beyond the scope of this book), the conversation about data ethics is at an interesting crossroads. Ethical principles are necessary, but not sufficient, and the form they take and how they are implemented will influence their effectiveness. Principles alone can effect little change unless embedded in policies and frameworks, and will always reflect the risk appetite of non-executive directors on a company board. A whole-of-business approach is required and this may strain many large organizations siloed in their functions, and challenge small ones not yet at this level of maturity. Either way, this is a new area of work that is not yet well enough understood and is rarely seen as a core part of a business.

As I read through this book, I noticed that one of its most prominent themes is the creation of a common language. When I present conference papers on similar topics, the looks of surprise in the audience and the frequency with which we talk past each other is a sign that we do not yet have a common understanding of ethics, nor do we have responsible building or use of technological systems. Just recently, I was told that someone 'did' responsible artificial intelligence, 'because women are involved'. These singular perspectives on data ethics are problematic and not indicative of the broader data governance roles emerging in Europe, roles that will become only more prevalent. This book is a valuable contribution to a joint conversation, adding to a generalist knowledge, understanding and lingo that we desperately need. As Statler and Waldorf on *The Muppets* would say, 'More, more!'.

References

Arendt, H. (1951) *The Origins of Totalitarianism*, Schocken Books, New York.

Byrne, E. (2021) 'Volkswagen loses High Court bid to challenge $125 million penalty for misleading customers', *ABC News*, available at https://www.abc.net.au/news/2021-11-12/volkswagen-diesel-emissions-penalty-high-court/100614788, accessed October 2023.

Kaminski, M. (2023) 'The developing law of AI regulation: a turn to risk regulation', *Lawfare*, 21 April, available at https://www.lawfaremedia.org/article/the-developing-law-of-ai-regulation-a-turn-to-risk-regulation, accessed October 2023.

Leins, K. (2023) 'Dr Kobi Leins (GAICD)'s post', available at https://www.linkedin.com/feed/update/urn:li:activity:7041248935293702144/, accessed October 2023.

Melbourne Social Equity Institute (n.d.) 'Why do they do that? An investigation into the perpetrator perspective of technology facilitated abuse', University of Melbourne, available at https://socialequity.unimelb.edu.au/projects/why-do-they-do-that, accessed October 2023.

Rosenblatt, J. (2023) 'Zuckerberg, TikTok "knew social media apps could harm teens"', *Sydney Morning Herald*, 14 March, available at https://www.smh.com.au/world/north-america/zuckerberg-tiktok-were-warned-about-social-media-addiction-in-teens-filing-shows-20230314-p5crt5.html, accessed October 2023.

Schiffer, Z. and Newton, C. (2023) 'Microsoft lays off team that taught employees how to make AI tools responsibly', *Verge*, 14 March, available at https://www.theverge.com/2023/3/13/23638823/microsoft-ethics-society-team-responsible-ai-layoffs, accessed October 2023.

Tangalakis-Lippert, K. (2023) 'Police are prosecuting abortion seekers using their digital data – and Facebook and Google help them do it', *Insider*, 5 March, available at https://www.businessinsider.com/police-getting-help-social-media-to-prosecute-people-seeking-abortions-2023-2?r=US&IR=T, accessed October 2023.

Tranter, K. and Bikundo, E. (2018) 'Carl Schmitt's *Die Buribunken*: a troubling critique of informational existence 100 years on', 16 July, *Griffith News*, available at https://news.griffith.edu.

au/2018/07/16/carl-schmitts-die-buribunken-a-troubling-critique-of-informational-existence-100-years-on/, accessed September 2023.

Winner, L. (2020) *The Whale and the Reactor: A Search for Limits in an Age of High Technology*, University of Chicago Press, Chicago.

Xiang, C. (2023) '"Horribly unethical": startup experimented on suicidal teens on social media with Chatbot', *Vice*, 7 March, available at https://www.vice.com/en/article/5d9m3a/horribly-unethical-startup-experimented-on-suicidal-teens-on-facebook-tumblr-with-chatbot, accessed October 2023.

Kobi Leins
King's College London
kobi.leins@kcl.ac.uk