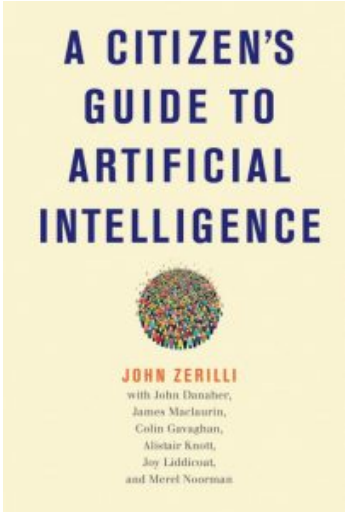
Book Review: A Citizen's Guide to Artificial Intelligence by John Zerilli, John Danaher, James Maclaurin, Colin Gavaghan, Alistair Knott, Joy Liddicoat and Merel Noorman

In **A Citizen's Guide to Artificial Intelligence***,* **John Zerilli***,* **John Danaher***,* **James Maclaurin***,* **Colin Gavaghan***,* **Alistair Knott***,* **Joy Liddicoat** *and* **Merel Noorman** *offer an overview of the moral, political, legal and economic implications of artificial intelligence (AI). Exemplary in the clarity of its explanations, the book provides an excellent foundation for considering the issues raised by the integration of AI into our societies, writes* **Karl Reimer***.*

This review originally appeared on LSE Review of Books. If you would like to contribute to the series, please contact the managing editor of LSE Review of Books, Dr Rosemary Deller, at lsereviewofbooks@lse.ac.uk

---

A Citizen's Guide to Artificial Intelligence. **John Zerilli, John Danaher, James Maclaurin, Colin Gavaghan, Alistair Knott, Joy Liddicoat and Merel Noorman. MIT Press. 2021.**



A Citizen's Guide to Artificial Intelligence is a text that ought to be read widely. The book's subject matter is highly relevant and it provokes many probing questions that deserve further consideration on the part of the reader and broader society.

The book itself covers a multitude of topics, ranging from 'What is Artificial Intelligence?', where the science behind Artificial Intelligence (AI) is described, to Deep Learning, machine learning, neural networks and other material. Later in the text, 'Algorithms in Government' and 'Oversight and Regulation' consider the integration of AI into daily life. Given space constraints, I will focus on two particular chapters in closer detail: 'Transparency' and 'Responsibility and Liability'.

Authors John Zerilli et al begin the 'Transparency' chapter with an anecdote suggesting one shouldn't trust technologies unless one has a way to investigate them (22). From this, they draw the relevant question: 'what exactly does it mean for a system to be transparent?' The remainder of the chapter is dedicated to expanding upon the various facets of transparency including responsibility, accountability, accessibility and inspectability.

Zerilli et al explain the legal concepts behind the right that an individual has to an explanation. They then contrast this against algorithmic systems that fail to provide reasonable explanations for their actions. Further, such algorithmic systems cannot be 'appealed' in their decisions as is the case in traditional legal cases (28). This then provokes the question of what explanations have been demanded of AI systems before the authors emphasise the relevant topic of standard-setting for AI. A given example of such a standard is the European Union's General Data Protection Regulation (GDPR) (30).

Date originally posted: 2021-10-16

Permalink: https://blogs.lse.ac.uk/impactofsocialsciences/2021/10/16/book-review-a-citizens-guide-to-artificial-intelligence-by-john-zerilli-john-danaher-james-maclaurin-colin-gavaghan-alistair-knott-joy-liddicoat-and-merel-noorman/

Blog homepage: https://blogs.lse.ac.uk/impactofsocialsciences/

The 'Responsibility and Liability' chapter is similar in its approach, which is evidence of the overall readability of the text. In this chapter, the leading questions are: 'do we want machines to be held responsible for decisions?' and 'can machines be responsible?' Drawing upon the work of jurisprudence scholar H.L.A. Hart, a hypothetical scenario of a drunk sea captain is given to highlight the complexity of responsibility (62). In the scenario, a fictional sea captain is responsible for the safety of his passengers and crew. However, he becomes drunk every evening of his voyage and at one point the ship is lost amidst a storm with no survivors. Because of his drunken state at the time of the accident, the sea captain is considered negligent and criminally responsible for the loss of life in legal proceedings following the accident. The important observation is that responsibility is a complex notion, which can refer to causal contribution as well as 'the obligations and duties that come with the professional role of a sea captain' (62). Just as responsibility is complex for the sea captain, so too is it complex for AI machines.

Zerilli et al note that responsibility comes in the form of moral responsibility – generally attached to individuals – and legal responsibility – attached to individuals as well as corporate entities such as Google or Facebook. Between AI technologies and human beings, there seems to be a 'responsibility gap' whereby it is difficult to hold human individuals responsible for actions (71). For example, one can consider how a multitude of programmers combine their efforts to create a singular automated driving system – one can argue it would be unreasonable to place responsibility onto an individual programmer. Toward the conclusion of the chapter, Zerilli et al also probe the possibility of 'morally and legally responsible AI', which indeed deserves consideration given the complexity of AI issues (76).

From a discussion of these two chapters, one can get a sense of the style Zerilli et al have used across the book. It is a well grounded and objective work. The text goes to the point of the actual issues that academics are working on at this moment. For this, the text is insightful and prescient: it neither presents the topic of AI as science fiction nor is it dismissive of the capabilities of AI. The text is also accessible to a wide audience. The authors come from legal and philosophical backgrounds yet are able to describe the complexities of AI systems with subtle nuance. The first chapter where the variety of AI systems is explained exemplifies this.

Date originally posted: 2021-10-16

Permalink: https://blogs.lse.ac.uk/impactofsocialsciences/2021/10/16/book-review-a-citizens-guide-to-artificial-intelligence-by-john-zerilli-john-danaher-james-maclaurin-colin-gavaghan-alistair-knott-joy-liddicoat-and-merel-noorman/

Blog homepage: https://blogs.lse.ac.uk/impactofsocialsciences/

The criticism I have of the text is that it is uneven in the quality of its content as well as its scope and breadth. Consider the brief introduction, 'Algorithms in Government'. Zerilli et al rightly observe a trade-off between legitimacy and efficiency (130), whereby policy agencies become removed from the citizenry who elected them when they are given increased discretionary unelected power. An example of this is the controversy surrounding UK GCSEs and A Levels based on a systematic calibration conducted by education regulator Ofqual. However, in this instance, as in others, there is little nuance to the argument made, which generally identifies a problem relating to AI, offers case studies showing this and then concludes that we need to think carefully about AI in society.

The 'Algorithms in Government' chapter could be extended to discuss further issues, such as whether an AI system can itself wield any legitimate authority. Further, the chapter on 'Control' could be summarised as an introduction to what scholars call the 'control problem' – when and whether humans ought to delegate effective control to AI systems. It is a reflection of the 'thin spread' of the book that this one problem has an entire chapter for its discussion as opposed to being integrated into questions of responsibility and legitimacy.

A Citizen's Guide to Artificial Intelligence is nonetheless a text that deserves to be read widely. It offers a sufficiently broad overview of the expansive literature on AI. It's a book that one could recommend to any individual without feeling guilty about sharing an overly complex topic. Zerilli et al are exemplary in the clarity of their explanations of AI and its influence on society.

The takeaway is this – AI as we conceive of it is already integrated in our society through the algorithms and automated decisions carried out by policymakers and corporate entities. As such, a deep consideration of these issues is necessary. The work put forward by Zerilli et al is an excellent foundation to this end.

Note: This review gives the views of the author, and not the position of the LSE Impact Blog, or of the London School of Economics.

Image Credit: Lukas via Unsplash.

Date originally posted: 2021-10-16

Permalink: https://blogs.lse.ac.uk/impactofsocialsciences/2021/10/16/book-review-a-citizens-guide-to-artificial-intelligence-by-john-zerilli-john-danaher-james-maclaurin-colin-gavaghan-alistair-knott-joy-liddicoat-and-merel-noorman/

Blog homepage: https://blogs.lse.ac.uk/impactofsocialsciences/