Fathom book club: Kearns, Lee and Pasquale

22 June 2022

Andrew Harris

Generally, Fathom's research notes aim to share our current findings for you to read. This time, however, we thought we'd mix things up – sharing our thoughts about our own reading.

Since Fathom held its first book club meeting a few months ago, a range of books have been in the limelight, prompting vigorous debate about the design of algorithms, the growth of AI and how to prevent artificial intelligence and robots from causing harm to human beings. Here are the latest books we have read, and what different Fathom economists have taken from each discussion.

The Ethical Algorithm – Michael Kearns

This interesting book describes how regulation could influence the way AI evolves: e.g., in protecting privacy, by randomising private data while preserving all the information at the aggregate level. So inferences like 'If you are a white male aged 56 in the UK you are likely to enjoy watching cricket' can still be drawn; but nobody can know whether Erik Britton fits that description (he does) unless I tell them so.

This is timely work, which gives examples of how easy it is for algorithms to inflict harm even when no harm was intended: even where the authors took steps to avoid harm. For example, it is well known that algorithmic ways of making decisions – such as whether to lend a sum of money to an individual – can codify latent patterns of discrimination, and even amplify those patterns. And that can happen even when the algorithms are coded to avoid any fields where discrimination tends to exist, like racial identity, or gender; this is because other, supposedly less sensitive fields, such as income group, postcode and job description, themselves embed the same patterns of discrimination. What fields are free from those patterns of discrimination? None. Which means the design of the algorithm must be adapted so that its objectives include not having any discriminatory patterns in its decisions. Which means it must have access to the controversial fields (race, gender etc) after all, simply to train itself to avoid discrimination.

I can see how with painstaking care we can make an ethical algorithm. What I can't see is how those ethical standards can be widely adopted. If there's a market for doing things the simpler and much worse way, that market will surely be filled. (Reviewed by Erik Britton, Managing Director/CEO)

AI Superpowers: China, Silicon Valley and the New World Order – Kai-Fu Lee

With AI likely to be the driving force behind the next wave of technological progress, Kai-Fu Lee takes us on a tour of the world's two AI superpowers. He argues that, while the US possesses the world's leading researchers, China's rapidly growing tech sector has access to better data. We concur with these points but are less convinced of the author's argument that the country with the data advantage will always win out. We are currently in the midst of a major consultancy project addressing exactly these questions. Conclusions will be revealed in due course… (Reviewed by Andrew Harris, Deputy Chief Economist)

New Laws of Robotics – Frank Pasquale

News that a Google engineer was recently put on leave after becoming convinced that his artificial intelligence (AI) was sentient has revived interest in the potential for increasingly humanlike AI. Whether you think that such an outcome is utopian or dystopian partly depends on which sci-fi movies you have watched, but it is also up to us — at least, that is the message from *New Laws of Robotics*, a 2020 book by Frank Pasquale, a professor at Brooklyn Law School.

In the book, Pasquale updates Isaac Asimov's Three Laws of Robotics,[1] offering four new rules to guide our interaction with machines and AI. There is a lot to like in his suggestions: robots should complement professionals, not replace them; AI should not counterfeit humanity; AI should not intensify zero-sum arms races; and AI must always indicate their human owner or creator, to ensure accountability. If achieved, these rules would certainly lessen anxiety about potentially harmful unintended consequences from rapid technological change.

I was left wondering however whether Pasquale's proposals were practical. The world has found it difficult to reach international agreement on its response to climate change. As with climate, there would be the potential for countries to 'free-ride' on any protocol on AI and robotics. The consequences look particularly fraught, considering that it is democracies in Europe and North America that would be most likely to implement more stringent regulation — leaving them at risk of being left behind by competitors in an increasingly multi-polar world. Such an outcome may be even scarier than sentient AI. (Reviewed by Kevin Loane, Head of Fathom USA)

COMING UP: next time in Fathom's book club we will be reading about husbanding the planet's resources, life on the sheep farm and the extraordinary world of fungi.

Further reading

Fathom book club: Copeland, Kahneman and Rozelle

Thank Fathom It's Friday: Big Data, or Big Brother?

---

1. Not to injure humans, to obey orders given by humans, and to protect itself as long as that does not violate laws one or two

Fathom Consulting
47 Bevenden Street
London
N1 6BH
Tel: +44 (0)20 7796 9561

Contact information
andrew.harris@fathom-consulting.com
www.fathom-consulting.com

Regulatory Disclaimer

FFC LIMITED and all of its affiliates (henceforth FFC) do not conduct "investment research" as defined in the FCA Conduct of Business Sourcebook (COBS) section 12 nor do they provide "advice about securities" as defined in the Regulation of Investment Advisors by the U.S. SEC. FFC is not regulated by the SEC or by the FCA or by any other regulatory body.

This research report has not been prepared in accordance with legal requirements designed to promote the independence of investment research and is not subject to any prohibition on dealing ahead of the dissemination of investment research. Nonetheless, FFC has an internal policy that prohibits "front-running" and that is designed to minimize the risk of receiving or misusing confidential or potentially material non-public information.

The views and conclusions expressed here may be changed without notice. FFC, its partners and employees make no representation about the completeness or accuracy of the data, calculations, information or opinions contained in this report. This report may not be copied, redistributed or reproduced in part or whole without FFC's express permission.

Information contained in this report or relied upon in its construction may previously have been disclosed under a consulting agreement with one or more clients. The prices of securities referred to in the report may rise or fall and past performance and forecasts should not be treated as a reliable indicator of future performance or results. This report is not directed to you if FFC is barred from doing business in your jurisdiction. Nor is it an offer or solicitation to buy or sell securities.

Analyst Certification

I Andrew Harris, the lead analyst, certify that the views expressed herein are mine and are clear, fair and not misleading at the time of publication. They have not been influenced by any relationship, either a personal relationship of mine or a relationship of the firm, to any entity described or referred to herein nor to any client of FFC nor has any inducement been received in relation to those views.

I further certify that in the preparation and publication of this report I have at all times followed all relevant FFC compliance protocols including those reasonably seeking to prevent the receipt or misuse of material non-public information.