

REVIEW ESSAY

There are No Facts: Attentive Algorithms, Extractive Data Practices, and the Quantification of Everyday Life Mark Shepard (2022), 296pp., \$28 hardback, MIT Press, Cambridge MA, ISBN 9780262047470

Introduction

Mark Shepard's book, *There are No Facts*, is a refreshing interdisciplinary work that explores the captivating crossroads of philosophy, technology and urban studies. With a robust methodological foundation, carefully selected and diverse case studies, this book explores the complex world of algorithms and machine learning, identifying prominent examples of social problems in the post-truth era, where elements of data monetization can be found in almost every aspect of our lives. In the data analytics society, individuals who consent to data collection by companies inadvertently allow them to gain knowledge about others. These harms go beyond the invasion of privacy and require new theoretical concepts to address the social costs (Pałka, 2020). This work acknowledges also the concepts introduced by Shoshana Zuboff, such as surveillance capitalism, which is regarded as a new economic order that recognizes human experience as a freely available resource (Zuboff, 2019). It expertly guides the reader through the key stages identified by the author, shedding light on why objective facts appear to be overshadowed by opinions and emotions in the contemporary world. The book is divided into two parts: the first is focused on 'practices' (Chapters 1–3), which build theoretical framework for the second part (Chapters 4–9), which provides the reader with interesting case studies.

Practices

In Chapter 1, 'Alternative facts', the author starts with the description of Donald Trump's inauguration in 2017, continues with the description of the deepfake of Barack Obama in 2018 and concludes with the presentation of notions of what is 'real' through the history of visual culture. The main aim of this chapter is to show how difficult it is to distinguish the truth from the 'alternative facts' in the post-truth era. And – in fact – this is just the beginning. The increasing size and sophistication of large language models, such as ChatGPT from OpenAI used for creating deepfakes, coupled with the potential rise of open source generative AI text-to-image models such as Midjourney, may enable further manipulation by malicious users (Edwards, 2022). While platforms may attempt to remove deepfake content, their success is uncertain, and the challenge of manually detecting deepfakes is expected to escalate, as demonstrated by research showing people's significant difficulty in distinguishing deepfakes from authentic videos, even in professional productions (Lewis *et al.*, 2022). In this area, China seems to be a global leader, especially after the announcement of Internet Information Service Deep Synthesis Management Provisions, which took effect in January 2023 (Cyberspace Administration of China, 2022). These provisions in Article 16 require a clear label of the synthesized content. The technological challenge is posed by the fact that once content is generated, it can be dissociated from the platform of its creation and distributed autonomously, released from the control of its original source (Hine and Floridi, 2022, p.609).

In Chapter 2, 'Ground fictions', Shepard provides interesting examples of cartography serving as a means of societal governance, prioritizing the creation of a shared imagination over

absolute accuracy. The governance, growth and decline of empires were influenced by the impact of maps which, in turn, were shaped by the empires themselves, determining their content, material representation and intended audience (Simpson, 2021). In this chapter, the author utilizes the example of the fictitious Sandy Island, initially charted by James Cook in 1776, which continued to be depicted on subsequent maps until the early twentieth century, to highlight the pervasive presence of falsehoods in our world that often go unnoticed or are unwittingly embraced (pp.65–6). Maps serve as purposeful narratives, which often exhibit inherent biases, partial perspectives and selective portrayals (Short, 2003, p.24).

In ‘The data blasé’ (Chapter 3), Shepard uses George Simmel’s blasé attitude (Simmel, 1903) to propose the novel approach of a ‘data blasé attitude’, according to which people are collectively dependent on data and network services (p.75). What is fascinating in this vision is that even after Edward Snowden’s disclosure in 2013 (Lyon, 2015), fears of intrusive surveillance are still quite limited within Western data-driven society (p.81). While China is mentioned several times in Shepard’s book, Chapter 3 pays particular attention to Chinese development of its social credit system as an example of the reputation economy as one of the elements of the data blasé attitude (pp.77–9). I really appreciate Shepard’s efforts to debunk the mistaken Orwellian myth propagated by some Western journalists, which have falsely portrayed the social credit system as a complete and coherent surveillance apparatus or as a national rating and moralizing mechanism (Knight, 2021). Furthermore, I appreciate his acknowledgement that the fears of Western journalists are not entirely unfounded, but rather stem from a wrong assumption that both public and private initiatives, such as Alibaba’s Sesame Credit system and limited-scale trials in specific cities and provinces, are part of the same system (pp.77–8). At this point it is worth noting that, with only an initial draft introduced in November 2022, there is still not fully developed national legal framework for the social credit system in the PRC (Von Blomberg, 2020; Knight, 2023). It is possible that the social credit system in China will continue to prioritize financial creditworthiness and regulatory compliance while maintaining minimal inclusion of individual behaviour, political preferences or social interactions and requiring clear official document-based regulations (Brussee, 2023, p.190). China’s example still serves as a commendable illustration of utilizing reputation as a form of currency.

Contexts

The second part of the book provides a comprehensive illustration of the practices presented in the first part. By showcasing real-life cases, the second part reinforces the relevance and efficacy of the aforementioned practices in various contexts. In Chapter 4, ‘Artificial cohabitants’, the author explores the infiltration of our daily life by such virtual assistant devices as Alexa from Amazon or Siri from Apple, which are constantly reconfiguring the boundaries between public and private spaces (p.102). Shepard mentions the problem of feminization of AI virtual personal assistants, including Alexa, Cortana and Siri, which exhibit a distinct female gender through their names, voices and characters (pp.101–2). This raises concerns about the perpetuation of societal norms that portray women as submissive and subordinate to men (Loideain and Adams, 2020). The emergence of digital representations of deceased individuals, commonly referred to as ‘ghostbots’, is the next step and another problem with various potential societal and individual harms concerning privacy, property, personal data and reputation (Harbinja *et al.*, 2023).

‘Spurious correlations’ are explored in Chapter 5, where Shepard examines shopping behaviour, focusing on the emergence of cashier-less shopping. To illustrate this concept, the author makes a case study of AmazonGo shops. A disruptive innovation, the cashier-less store revolutionizes the retail experience by eliminating the need for human or automatic cashiers, enabling customers to enjoy a seamless shopping experience without wasting time in long queues or self-checkout. The system tracks their selections and preferences effortlessly (Gazzola *et al.*, 2022, p.5). These shops utilize a wide array of cameras, and Shepard connects this technology with the application of Bayesian statistical methods (p.115). The aim is to demonstrate how retailers can leverage

large volumes of data to make predictions about the future purchasing decisions of their customers. Furthermore, Shepard incorporates Shoshana Zuboff's concept of a behavioural futures market. This concept highlights how companies can anticipate and make informed bets on the future behaviour of consumers (Zuboff, 2019). By combining Zuboff's idea with the discussions on cashier-less shopping and the utilization of vast datasets, the author presents a comprehensive perspective on how retailers can utilize advanced technologies and predictive analytics to forecast and influence consumer behaviour. It is worth noting that this technology is also used extensively by Żabka, the Polish convenience shops chain, which is poised to become the world's largest operator of autonomous stores with over 50 locations (including one in the Tesla Gigafactory in Berlin), solidifying its leading position in Europe's autonomous retail market (Hackett, 2023).

In Chapter 6, 'From tools to environments', Shepard proves that big data and machine learning are not less biased than the human bias inherent in the data used to train them (p.122). In this chapter, he examines environments to highlight the changing bias in the design of urban systems and the resulting epistemological implications. He analyses two different approaches to urban research that utilize image analysis techniques, comparing small data studies with big data analysis. Using the Hudson Yards development in Manhattan as an example, Shepard examines how the measurement of neighbourhood activity involves its residents in the data collection process (pp.133–5). The application of concepts and techniques from behavioural economics to an entire community presents a mode of governance on an urban scale. In *A City is Not a Computer*, Shannon Mattern (2021) contends that computational models of urbanism limit our understanding of cities by overshadowing the importance of local intelligence and knowledge institutions, which, she argues, are crucial for providing a necessary supplement and correction to prevailing algorithmic approaches. Shepard concludes the chapter by questioning how the quantification of urban life through smart city initiatives is not only transforming our understanding of cities, but also influencing our perception of their inhabitants, which might be inherently biased.

'The right to the (wrong) city' follows. Here Mark Shepard looks at the impact of attentive algorithms and extractive data practices on urban districts. He highlights the similarities between these issues and the ones previously discussed in relation to data extraction, epistemic fragmentation and subject positioning in the home, urban minimarkets and neighbourhoods. Shepard emphasizes the need to examine how these issues affect governance and governmentality on an urban scale. The smart city combines traditional and modern infrastructure, leveraging the internet of things and information and communication technology to collect and process data for resource management and advanced services, with the aim of providing convenient and reliable services that enhance quality of life at a reduced cost, covering areas such as power grids, traffic waste management and healthcare (Rani *et al.*, 2022). Shepard discusses the controversies regarding data privacy as well as much broader problems arising from the emergence of urban-scale data determinism and the inclusivity or exclusivity of individuals in the described process. To illustrate these concerns, Shepard discusses the problems with development of a smart waterfront in Toronto. The Quayside in Toronto illustrates the ambition to create a sustainable, resilient and innovative urban community while raising questions about the reconfiguration of urban governance and the inclusivity of stakeholders in shaping urban life.

In Chapter 8, 'The ruse and the exploit', Mark Shepard explores the application of ruse, deception and exploitation within the context of political campaigns to create divisions within national electorates. Taking the role of the 'election management' company, Cambridge Analytica, in the 2016 United States presidential election and the Brexit referendum as a starting point, he traces how concepts and techniques borrowed from behavioural economics by marketing and advertising have been introduced into contemporary political campaigns, from psychometric micro-targeting to the creation and dissemination of synthetic media through social media platforms. Shepard concludes by examining the implications of this deviation not only for electoral politics, but also for broader matters such as public space, representative democracies and the epistemology of uncertainty. Furthermore, the new wave of advanced generative AI models brings forth

the potential risk of learning from datasets comprising text and images sourced from the internet, which may inherently contain biases and prejudices, leading to the perpetuation and amplification of social inequalities, reinforcement of stereotypes and the emergence of new forms of discrimination in the generated content (Tredinnick and Laybats, 2023).

‘Pandemic exceptionalism’ sees the author delving into the various examples that illustrate how the global community responded differently to the covid-19 pandemic in various countries. The chapter sheds light on the role of data in this process, with a particular focus on the controversies surrounding data collection and data visualization. Shepard also explores the potential for different types of data manipulation (p.179). This is especially interesting as a visible relationship exists between users’ trust in the government and their perception of privacy protection (Huang *et al.*, 2022). Throughout the chapter, Shepard highlights the diverse approaches taken by governments and institutions worldwide in dealing with the pandemic. These approaches vary significantly, ranging from strict lockdown measures and rigorous testing to more relaxed strategies that prioritize economic stability. By examining these variations, Shepard provides a nuanced understanding of how different nations have navigated the challenges posed by mapping pandemic data and how the pandemic space was represented. For example, the Chinese government implemented a mandatory smartphone application that tracks residents’ movements and assigns them colour codes (red, yellow or green) based on perceived public health risk, regulating access to such public spaces as subways and malls (Ram and Gray, 2020, p.2). However, concerns arise as the methodology for determining these colour codes remains opaque, and the app’s sharing of information with law enforcement sets a precedent for potential long-term automated social control measures beyond the pandemic, a prospect that is promoted by both officials and technology firms (Chen *et al.*, 2022).

According to researchers, the health code system in China, along with the social credit system, is considered a data-driven experiment that challenges the multi-stakeholder model of digital governance (Yu, 2022). China’s evolving data protection regime, with negotiable data classifications and vague categorizations, along with widespread consent forms, provides Chinese technology firms with ample room to utilize amassed data for developing new AI products, potentially strengthening their position compared with foreign firms lacking similar access (Parasol, 2021). Various European countries have developed tracing apps with different technological approaches, including such geolocation-based apps as the Spanish CoronaMadrid and StopCovid19, as well as such Bluetooth-based apps as the Polish ProteGo, the Austrian Stopp-CoronaApp and the French StopCovid (Bradford *et al.*, 2020, p.3).

One key aspect that Shepard emphasizes is the importance of data in informing decision-making during a global health crisis. Data collection plays a vital role in tracking the spread of the virus, understanding its impact on populations and evaluating the effectiveness of various containment measures. However, Shepard also acknowledges the controversies surrounding data collection methods and the potential biases that can arise from them, mentioning examples from Florida in 2020, undercounting deaths in nursing homes in New York and the false visualization of a downward trend in covid-19 cases in Georgia (p.181). In addition to data collection, Shepard explores the significance of data visualization in shaping public perception and influencing policy responses. Visual representations of data, such as charts, graphs and maps, could have a profound impact on how people interpreted and responded to the pandemic. The author discusses how different visualizations could highlight or downplay certain aspects of the crisis, leading to varying levels of public awareness and government action. Furthermore, Shepard delves into the potential for data manipulation, raising concerns about how data can be selectively presented or distorted to support particular narratives or agendas. This manipulation can occur at various stages, including data collection, analysis and presentation. By shedding light on these possibilities, Shepard prompts readers to critically evaluate the information they encounter and to be wary of potential biases or inaccuracies. Overall, this chapter serves as a thought-provoking exploration of the complexities inherent in global responses to the pandemic, urging readers critically to analyze the information they encounter and to recognize the influence of data in shaping our understanding of the crisis.

Towards an even more terrifying future

In the final chapter, ‘Coda’, Mark Shepard raises the question of our ability to take action and collectively engage in the complex interrelationships among people, data, code, space, knowledge and power. Today, public space is no longer defined solely by physical geography and instead, as the chapters in this book argue, the public realm is increasingly constituted by spaces of surveillance rather than spaces of appearance. The challenge we face is to articulate new forms of appearance within contemporary public spaces where we come together, horizontally spanning and countering the divisive vertical divisions of surveillance spaces. This would support an ongoing struggle to formulate more empowering and enduring collective identities. In conclusion, the most important issue the author unearths is: how we navigate and negotiate this epistemic uncertainty and learn to live with accompanying doubt will likely shape the futures we aspire to share and those we must strive to prevent.

I believe this question holds greater relevance in the current context, particularly when considering the recent emergence of highly advanced pretrained generative models such as OpenAI’s GPT-4, the Chinese ERNIE Bot from Baidu and Google’s Bard, which have demonstrated remarkably promising performance and surpassed previous levels of accuracy (Xiao *et al.*, 2021, p.79; Tredinnick and Laybats, 2023). The increasing disparity between the rapid pace of technological advancements and the lag in regulatory measures inevitably leads to outdated or overlooked attempts at regulating these new technological developments (Marchant, 2011, p.19). Generative AI, with its ability to produce convincing and sophisticated content, presents a significant concern as it could amplify the spread of disinformation during elections, economic crises and wars anywhere in the world (Tredinnick and Laybats, 2023).

Conclusions

I highly recommend this book to anyone, but especially to legal scholars and computer scientists. Shepard’s unique perspective sheds light on how we perceive attentive algorithms, data extraction and monetization in the dawn of the era of generative AI and deep synthesis. Overall, this book serves as a perfect introduction to better understanding of how data could be monetized, corrupted and misused. Mark Shepard’s interdisciplinary background as an architect and artist provides extremely interesting illustrations of the problem using examples from the methodology of philosophy, urban studies and art. Legal scholars will benefit from Shepard’s insights into the legal implications surrounding data monetization, corruption and misuse, which could lead to more innovative suggestions for regulation. Computer scientists, on the other hand, will gain valuable insights into the ethical considerations and technical aspects of developing and implementing attentive algorithms. By delving into the pages of this book, readers will be equipped with a more comprehensive understanding of the complex interplay among data, algorithms and societal impact. Shepard’s unique perspective encourages readers to evaluate critically the potential consequences and implications of these technologies, fostering a proactive approach towards responsible and ethical data practices. In conclusion, this book stands as a valuable resource for anyone seeking a fresh perspective on how technological development relates to truth. Shepard’s interdisciplinary background provides captivating illustrations that vividly portray the challenges of disinformation and the immense power wielded by data processing technologies.

References

- Bradford, L., Aboy, M. and Liddell, K. (2020) ‘COVID-19 contact tracing apps: a stress test for privacy, the GDPR, and data protection regimes’, *Journal of Law and the Biosciences*, 7, 1, pp.1–21.
- Brussee, V. (2023) *Social Credit: The Warring States of China’s Emerging Data Empire*, Palgrave Macmillan, Singapore.

- Chen, W., Huang, G. and Hu, A. (2022) 'Red, yellow, green or golden: the post-pandemic future of China's health code apps', *Information Communication and Society*, 25, 11, pp.1–16.
- Cohen, J. (2019) *Between Truth and Power: The Legal Constructions of Informational Capitalism*, Oxford University Press, Oxford.
- Cyberspace Administration of China (2022) *Internet Information Service Deep Synthesis Management Provisions* (互联网信息服务深度合成管理规定), available at https://www.gov.cn/zhengce/zhengceku/2022-12/12/content_5731431.htm, accessed June 2023.
- Edwards, L. (2022) 'Deepfakes in the courts', *Counsel*, December, available at [counselmagazine.co.uk/articles/deepfakes-in-the-courts](https://www.counselmagazine.co.uk/articles/deepfakes-in-the-courts), accessed June 2023.
- Gazzola, P., Grechi, D., Martinelli, I. and Pezzetti, R. (2022) 'The innovation of the cashierless store: a preliminary analysis in Italy', *Sustainability*, 14, 4, pp.1–16.
- Hackett, A. (2023) 'Żabka – largest operator of autonomous stores in Europe – aims to overtake Amazon as world's biggest', *Notes from Poland*, 23 January, available at notesfrompoland.com/2023/01/23/polands-zabka-now-largest-operator-of-autonomous-stores-in-europe-aims-to-be-worlds-biggest/, accessed May 2023.
- Harbinja, E., Edwards, L. and McVey, M. (2023) 'Governing ghostbots', *Computer Law & Security Review*, 48, pp.1–12.
- Hine, E. and Floridi, L. (2022) 'New deepfake regulations in China are a tool for social stability, but at what cost?', *Nature Machine Intelligence*, 4, pp.608–10.
- Huang, G., Hu, A. and Chen, W. (2022) 'Privacy at risk? Understanding the perceived privacy protection of health code apps in China', *Big Data & Society*, 9, 2, pp.1–13.
- Knight, A. (2021) 'Technologies of risk and discipline in China's social credit system' in Creemers, R. and Trevaskes, S. (eds) *Law and the Party in China*, Cambridge University Press, Cambridge, pp.237–62.
- Knight, A. (2023) 'Basket case: reform and China's social credit law', *China Law and Society Review*, 6, 2, pp.181–210.
- Lewis, A. *et al.* (2022) 'Do content warnings help people spot a deepfake? Evidence from two experiments', Royal Society, available at royalsociety.org/-/media/policy/projects/online-information-environment/do-content-warnings-help-people-spot-a-deepfake.pdf, accessed September 2023.
- Loideain, N. and Adams, R. (2020) 'From Alexa to Siri and the GDPR: the gendering of Virtual Personal Assistants and the role of Data Protection Impact Assessments', *Computer Law & Security Review*, 36, pp.1–14.
- Lyon, D. (2015) *Surveillance After Snowden*, Polity Press, Cambridge
- Marchant, G. (2011) 'The growing gap between emerging technologies and the law' in Marchant, G., Allenby, B. and Herkert, J. (eds) *The Growing Gap between Emerging Technologies and Legal-Ethical Oversight. The Pacing Problem*, Springer, Heidelberg, pp.19–33.
- Mattern, S. (2021) *A City is Not a Computer: Other Urban Intelligences*, Princeton University Press, Princeton NJ.
- Pałka, P. (2020) 'Data management law for the 2020s: the lost origins and the new needs', *Buffalo Law Review*, 68, 2, pp.559–640.
- Parasol, M. (2021) *AI Development and the 'Fuzzy Logic' of Chinese Cyber Security and Data Laws*, Cambridge University Press, Cambridge.

- Ram, N. and Gray, D. (2020) 'Mass surveillance in the age of COVID-19', *Journal of Law and the Biosciences*, 7, 1, pp.1–17.
- Rani, S. *et al.* (2022) 'Security and privacy challenges in the deployment of cyber-physical systems in smart city applications: state-of-art work', *Materials Today: Proceedings*, 62, 7, pp.4671–6.
- Short, J. (2003) *The World through Maps: A History of Cartography*, Firefly Books, Toronto.
- Simmel, G. (1903) 'Die großstädte und das geistesleben' in Petermann, T. (ed.) *Die Großstadt: Vorträge und Aufsätze zur Städteausstellung*, Von Zahn & Jaensch, Dresden, pp.187–206.
- Simpson, T. (2021) 'Cartography and empire from early modernity to postmodernity', *The Routledge Handbook of Science and Empire*, Routledge, London, pp.21–34.
- Tredinnick, L. and Laybats, C. (2023) 'The dangers of generative artificial intelligence', *Business Information Review*, 40, 2, pp.46–8.
- Von Blomberg, M. (2020) 'The social credit system and China's rule of law' in Everling, O. (ed.) *Social Credit Rating. Reputation und Vertrauen Beurteilen*, SpringerGabler, Wiesbaden, pp.111–37.
- Xiao, C., Hu, X., Liu, Z., Tu, C. and Sun, M. (2021) 'Lawformer: a pre-trained language model for Chinese legal long documents', *AI Open*, 2, pp.79–84.
- Yu, H. (2022) 'Living in the era of codes: a reflection on China's healthcode system', *BioSocieties*, 1, pp.1–18.
- Zuboff, S. (2019) *The Age of Surveillance Capitalism*, Public Affairs, New York.

Igor Szpotakowski
Newcastle Law School,
Newcastle University
I.M.Szpotakowski2@newcastle.ac.uk