EVERYDAY DATA CULTURES

JEAN BURGESS, KATH ALBURY,
ANTHONY McCOSKER AND ROWAN WILKEN

Everyday Data Cultures

Everyday Data Cultures

Jean Burgess, Kath Albury,
Anthony McCosker and
Rowan Wilken

polity

Copyright © Jean Burgess, Kath Albury, Anthony McCosker and Rowan Wilken 2022

The right of Jean Burgess, Kath Albury, Anthony McCosker and Rowan Wilken to be identified as Authors of this Work has been asserted in accordance with the UK Copyright, Designs and Patents Act 1988.

First published in 2022 by Polity Press

Polity Press
65 Bridge Street
Cambridge CB2 1UR, UK

Polity Press
101 Station Landing
Suite 300
Medford, MA 02155, USA

All rights reserved. Except for the quotation of short passages for the purpose of criticism and review, no part of this publication may be reproduced, stored in a retrieval system or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the publisher.

ISBN-13: 978-1-5095-4755-5
ISBN-13: 978-1-5095-4756-2(pb)

A catalogue record for this book is available from the British Library.

Library of Congress Control Number: 2021951508

Typeset in 11.5 on 14pt Sabon
by Fakenham Prepress Solutions, Fakenham, Norfolk NR21 8NL
Printed and bound in Great Britain by CPI Group (UK) Ltd, Croydon

The publisher has used its best endeavours to ensure that the URLs for external websites referred to in this book are correct and active at the time of going to press. However, the publisher has no responsibility for the websites and can make no guarantee that a site will remain live or that the content is or will remain appropriate.

Every effort has been made to trace all copyright holders, but if any have been overlooked the publisher will be pleased to include any necessary credits in any subsequent reprint or edition.

For further information on Polity, visit our website:
politybooks.com

Contents

Acknowledgements vi

1. Introduction 1

2. The Everyday Data Cultures Framework 25

3. Everyday Data Intimacies 46

4. Everyday Data Literacies 83

5. Everyday Data Publics 115

6. Conclusion 144

Notes 157
References 160
Index 195

Acknowledgements

This project received support from the Australian Research Council Centre of Excellence for Automated Decision-Making and Society (ADM+S). The #spotify-wrapped Twitter data discussed in Chapters 3 and 5 was collected and processed with assistance from Betsy Alpert and the QUT Digital Observatory. The dating apps project discussed in Chapter 3 was supported by the Australian Research Council Linkage Project 'Safety, Risk and Wellbeing on Dating Apps' (LP160101687), in partnership with ACON Health and Family Planning NSW. Kath Albury and Anthony McCosker gratefully acknowledge their collaborators on that project, particularly Paul Byron, Teddy Cook, Christopher Dietzel, Tinonee Pym, Kane Race, Daniel Reeders, Doreen Salon, Son Vivienne and Jarrod Walshe. We would also like to thank the anonymous reviewers of the draft manuscript for their insightful, supportive and constructive feedback.

1
Introduction

A single parent is working from home as usual in Melbourne, Australia, connecting to the internet using one of the Google Wi-Fi points that are distributed strategically throughout her terrace house. She's prepared to trade off the privacy risks associated with the data these devices gather for the comfort she derives from their aesthetically pleasing sleek design, and for the convenience they provide compared to the complication and hassle of other home Wi-Fi set-ups, especially when renting. Meanwhile, two doors down, a man thumbs through search results and targeted ads on his phone as he researches the planned purchase of a new smart TV. He is struck by the low price of certain models, but feels uneasy, vaguely recalling reading somewhere that some models are only priced so competitively because the TV makers collect user data and sell it on to third parties (Gilbert, 2019).

In Madrid, a teenager opens Snapchat and scrolls through newly received snaps. Because his virtual presence on the Snap Map (his Actionmoji) is moving

in a constant direction at speed, the Snapchat app correctly assumes that he is seated on a train. Beside him, his friend, who has been scrolling TikTok continuously, sees a video pop up in her 'For You Page'. It seems out of place compared to the videos the algorithm usually selects for her, and features a popular influencer humorously suggesting she take a break, have a drink of water and perhaps go for a walk. Barely pausing to roll her eyes at the clumsy intervention, she scrolls on to the next video in her feed (Burke, 2020). In a Brazilian megacity, a young woman is about to head out on her bicycle, and needs to work out the safest and most terrain-friendly route to take. Because she doesn't entirely trust app-generated directional recommendations, she selects her preferred route based on a combination of saved Strava and pre-downloaded Google Maps data and her own experiential knowledge of the city (Pink et al., 2019, p. 179).

Around the same time, a Japanese TV show reveals that the person behind social media star @azusagakuyuki, an attractive young woman who poses beside her motorcycle, is Soya, a fifty-year-old man. Soya created his highly successful female alter ego using AI-enabled face-editing apps – marketed as 'fun' apps, but licensed to users on terms that enable companies to collect large amounts of personally identifying data. When asked about it, Soya explains that, while he began by just playing around with the app, 'it happened to turn out to be fairly pretty' – and so @azusagakuyuki was born. Encouraged by the 'likes' he got after posting the results, he said, 'I got carried away gradually as I tried to make it cuter' (BBC News, 2021). Nearby in South Korea, Seo-Hyun, a young woman interested in fashion and beauty, gives careful consideration to the forms of 'zero party data' – 'personal data a consumer

intentionally and proactively shares' with brands and marketers (Mitchell, 2019) – that she is prepared to share with clothing and cosmetics brands, tailoring this information in such a way as to maximise her chances of being rewarded with promotional give-aways of her favourite jeans and make-up.[1]

Outside a courthouse in Oakland, California, an activist taking part in a demonstration against police violence is confronted by a local officer, and so the protestor begins recording video of their interaction on his phone. The police officer retaliates by pulling his own phone out of his pocket. To the bemusement of the watching crowd, he opens Spotify and starts playing a track by mainstream pop artist Taylor Swift, assuming that YouTube's copyright enforcement algorithms (YouTube, 2021) will use audio data-matching to automatically detect and remove the protestor's video, preventing it from reaching a large audience (Cabanatuan, 2021; Schiffer and Robertson, 2021). Nevertheless, the video goes viral on Reddit and Twitter, provoking discussions and knowledge-sharing about audio editing techniques that can be used to work around the data logics of the major platforms' automated content-moderation techniques.[2]

These composite vignettes are a mix of factual, semi-fictionalised and fictionalised accounts, designed to give the reader (you) a way into the book's themes and ideas.[3] They have also fulfilled an important function for us as authors. In preparing this book, the creation of vignettes aided us in compiling and distilling what we have observed over the last decade or so of our own and colleagues' qualitative research in this area, and in isolating what we see as important about everyday data cultures. In this way, the above vignettes serve as 'presentation devices' (Ely et al., 1997, p. 74): they

introduce and synthesise themes that are central to the book and to which we return, in different ways and to varying degrees, in later chapters. Not only do these vignettes reinforce how data is often both collected from and targeted to us as we go about our day-to-day lives (the datafication of everyday life), they also provide glimpses of how we form intimate relationships with and through data (everyday data intimacies), how we develop skills and capacities to do things with data (everyday data literacies), and how everyday data practices play out in communities and in public (everyday data publics).

The turn to data

At the most basic level, our opening vignettes together paint a preliminary picture of what is often referred to as the 'datafication' of everyday life. Our diverse daily activities – from connecting to the internet at home to managing our daily commute – are not only increasingly dependent on data; these activities and experiences are also routinely converted into digital data via our use of mobile devices, apps and the sensor networks embedded in our environments. And, increasingly, this data is used for automation and algorithmic decision-making processes, the results of which are fed back into our everyday lives.

As part of the first stage of this transition to datafication, services supporting a wide range of mundane activities and essential services (including media consumption, banking, shopping, transport, interpersonal communication and dating) have been steadily migrating to mobile apps and platforms with which we engage via data-driven identities and personal

data profiles (as happens, for example, when we use Facebook to log in to a wide variety of other websites and apps). Meanwhile, an increasing number of data-collecting sensors and internet-connected objects are becoming embedded in personal devices, 'smart' homes, workplaces and public spaces. And digital media platforms integrate into their operations various forms of algorithmic curation and artificial intelligence, which feed on, learn from and act on this data, thereby shaping our access to information, our cultural experiences and our relationships.

Datafication represents a significant moment in the history of technology and society. For critical scholars, the concept of datafication expresses the capacity of 'commercial digital media [to] capture the details of activities that once eluded systematic forms of value extraction in order to turn them into information commodities' (Andrejevic, 2010, p. 90). Relatedly, datafication is thought to be transforming our identities as both consumers and citizens: the increased 'scope and sophistication' of data collection and processing has made datafication 'a cornerstone of contemporary forms of governance', enabling 'both corporate and state actors to profile, sort and categorize populations' (Hintz, Dencik and Wahl-Jorgensen, 2019, pp. 2–3). Mejias and Couldry (2019) put the situation in even more dramatic terms: for them, the term 'has quickly acquired an additional meaning: the wider transformation of human life so that its elements can be a continual source of data'.

But the datafication of everyday life, especially by state and corporate actors, is also a recent manifestation of longer-term trends. History is full of attempts to 'pin down' and fix as information the most difficult-to-grasp aspects of everyday life, precisely

because everyday life was thought to be at the heart of what makes societies tick, how they change – and, hence, how their populations can be 'nudged' in one direction rather than another. For example, both Ben Highmore (2015) and John Storey (2014) devote entire chapters of their respective books on everyday life to the Mass-Observation project – the grandly ambitious attempt, undertaken by British (middle-class) researchers in the 1930s, to seek understandings of the lived experience of (working-class) everyday life by gathering first-hand information (from the 'masses') about it. The project generated large amounts of material, filling thousands of boxes with 'accounts of nightmares; meticulously detailed records of drinking habits in Bolton pubs (timed to the second with a stopwatch); pages and pages of diary records; [and] thoughts on margarine' (Highmore, 2015, p. 75).

And of course, audiences have been calculated (measured, segmented and targeted) by the media industries for as long as they have existed. In an earlier period of media studies focused on the challenges of studying television, John Hartley argued that broadcast television's audiences were 'invisible fictions' who had no existence outside the methods brought to bear on them by critics, policy-makers and industry actors. As Hartley points out, audiences need 'constant hailing and guidance' from the industry in order to see themselves *as* audiences – a system for 'imagining the unimaginable; for controlling the uncontrollable' (1987, p. 136). We see echoes of Hartley's ideas in the way that the large internet platforms have constructed, measured, targeted and addressed (hailed) us – these days, not as audiences but as 'users' – over the past forty years, continuing on the trajectory of inventing data and metrics (from likes to retweets) that can stand in for audience practices that

might otherwise remain private and personal. In the post-broadcast era, where broadcast, print and internet media have converged, these processes of audience measurement and segmentation, and the metricisation of their attention, are intensified (Burgess and Baym, 2020; Livingstone, 2004, 2019).

In media studies, the turn to datafication is often discussed as part of the debate about a trajectory from mediation to mediatisation (for an excellent overview of the attendant debates, see Livingstone and Lunt, 2016). Since media are thought to not only represent but to help shape social realities, changes in media – understood both as communication technologies and as the symbolic representations transmitted by those technologies (Silverstone, 2005) – have always been connected to changes in society. The idea of a shift to 'mediatisation', though, is that the organising principles and values (the 'logics') of the media system begin to play a dominant role in society more generally (Couldry and Hepp, 2017). Building on Altheide and Snow's (1979) idea of 'media logic', José van Dijck and Thomas Poell (2013) discuss how characteristics of 'social media logic' – like data-driven popularity metrics – have come to influence far broader spheres of social and economic life. Most recently, some media theorists have proposed the idea of 'deep mediatisation' (Couldry and Hepp, 2017; Hepp, 2019), as a further intensification of these tendencies, and one particularly tied up with datafication and platforms (Andersen, 2018). Under the data logics of 'deep mediatisation', not only might a restaurant need to be visible on social media, the owners might even think about redesigning the space or plating menu items in ways that are specifically targeted to maximise engagement on Instagram, with

the visual data processing and algorithmic logics of the platform in mind (as the restaurant web-hosting company *Owner*'s 'Ultimate Guide to Instagram for Restaurant Professionals' outlines in forensic detail).[4]

Importantly, it is possible for people to be affected by these processes of datafication and 'deep mediatisation' even when they don't have much access to or active engagement with digital technologies, meaning that we can be subject to these systems but have limited agency within them. Digital inclusion (and exclusion) remains an issue in almost every country (see, for example, Thomas et al., 2018), and in the context of datafication and automation, its implications for social inequality are only increasing. A lack of universally affordable, meaningful access to telecommunications data is one major source of inequality (Moyo and Munoriyarwa, 2021), intersecting with other aspects of digital inclusion or exclusion such as digital skills and confidence (Park and Humphry, 2019). Digital inclusion has particularly acute implications for people with disability, while at the same time their experiences afford 'a rich and indispensable site and "test bed" for how societies can confront technology for better futures' (Goggin, Ellis and Hawkins, 2019).

Our opening vignettes also hint at one of our principal interests in this book: the ways people live with, work around, or resist these data operations. A core argument here is that people aren't always only subject *to* datafication. Instead – in certain ways, at certain times and within certain constraints – we are active agents and, sometimes, disrupters and resisters *of* datafication as well. If nothing else, everyday experiences of technology are messy – even the 'smartest' of smart technologies never really work quite as seamlessly as either the marketing hype or the panicked media

stories about them would have us believe. At a deeper level, everyday life is where the politics of digital transformation are worked through in practice, as people negotiate, wrangle, learn and struggle with or against data-intensive technologies, in the context of their own bodies, lives, communities and histories. We see these politics played out in different ways in the earlier vignettes: the Japanese biker's off-label use of a face-editing app; the young South Korean woman's canny manipulation of zero party data in the context of intensive ad targeting; and the cyclist's circumspection around the safe use of navigational apps in the city she knows so well.

'Everyday data cultures' is a cultural studies-based conceptual framework that can be used to explore how the broader digital transformations associated with datafication (or 'deep mediatisation') are playing out at ground level, and how the activities, thoughts and feelings of citizens, consumers and users of technologies play a part in these processes. In what follows, we discuss the many ways data is created, transformed and shared in and through people's daily activities. But in turn, people's everyday practices of and ideas about data influence and shape Big Tech's data logics, infrastructures and flows, adding friction and noise to their business and revenue models. This interplay between everyday cultures of use and business logics can be seen playing out in different ways in the earlier vignettes: in TikTok's health interventions and the user's eye-rolling resistance to them; and in police counter-tactics that seek, in the most mundane and off-hand manner, to weaponise automated platform responses to copyright infringement. These 'everyday data cultures', in all their variety, richness and political contradictions, are the focus of this book.

Critical data studies and beyond

In the current period of intensive activity and media coverage related to artificial intelligence (AI), algorithms and automation, there is a growing chorus of critics and scholars sounding the alarm. These voices articulate an increasing concern about the take-up and power of data-intensive, automated decision-making technologies, both in terms of their ubiquity and in terms of the new or intensified forms of inequality and injustice that can result from them (see, for example, Eubanks, 2018; Amoore, 2020).

Intervening in the previous period of hype around 'big data' in the 2010s, scholars critiqued the inflated claims for its analytical power and revolutionary potential (Puschmann and Burgess, 2014), the problems of bias in applications of AI, such as predictive policing (boyd and Crawford, 2012), and the way the apparent seamlessness of data operations obscures infrastructures of logistics and labour. This work was important in highlighting the gaps between public understandings of and industry hype around both the benefits and dangers of datafication.

Since that first wave, the field of critical data studies has emerged at the intersections of digital sociology, cultural studies and internet studies (Iliadis and Russo, 2016), and has made significant advances in theorising, and diagnosing the politics of, data and datafication (Cheney-Lippold, 2017). The collection, reuse and exploitation of personal data by both corporate and government organisations has provoked concerns about trust, privacy and surveillance, leading to calls for new data rights (Ruppert et al., 2017), improved data literacy (McCosker, 2017a; Fotopoulou, 2021) and shared principles for data ethics (Zwitter, 2014).

Data has a cultural politics, too. As Catherine D'Ignazio and Lauren F. Klein remind us in their book *Data Feminism* (2020), the politics of data lie not only in what it includes or leaves out, or even what it *does*, but in the very idea of data, with its connotations of objectivity, its evidentiary power, and the binary logics through which it is so often constructed and deployed. Prefiguring and accompanying concerns about the uneven benefits of AI and datafication, and the problems with fairness and bias that can result, important interventions have also been made that challenge the unthinking centring of whiteness in technoculture (Brock, 2020), the racial bias of much data-driven decision-making, and the real-world impacts of racist data discrimination (Benjamin, 2019; Noble, 2018). Such interventions have already made a noticeable impact on the public conversation in this area, and have provided the impetus for potentially significant change within the technology industry – see, for example, the acclaimed documentary film *Coded Bias* (Kantayya, 2020), which featured some of the most prominent critical social scientists and tech activists in the United States, including former Google employees.

This critical data studies conversation has recently and very noticeably gone mainstream, and in the process it has taken a distinctive turn towards what we call 'Big Critique', by which we mean writing marked by a sense of understandable urgency expressed in increasingly polemical terms, set up principally against the unprecedented power of the large technology companies – primarily US-based companies like Google, Facebook and Amazon. To a lesser extent in Anglophone scholarship, given its tendency to US-centricity, such concerns also include Chinese companies like Tencent and BytePlus.

Big Critique is often characterised by bold new concepts and ideas that describe large-scale, whole-of-society (or whole-of-planet) concerns. In *Atlas of AI*, for example, Kate Crawford (2021) addresses at planetary scale the profit-driven logics and damaging social and environmental impacts of AI and the data it runs on. Mark Andrejevic's *Automated Media* (2019a) argues that data-driven automation brings with it a deep paradigm shift, wherein logics of prediction rather than representation, and the frameless 'fantasy of total information capture' (Andrejevic, 2019b), threaten not only to further entrench surveillance but also to close down the space of possibility for political action.

Nick Couldry and Ulises A. Mejias's *The Costs of Connection* (2019) provides an expansive theory of datafication in terms of colonialism, and the threat to human autonomy and social life itself that, they argue, the 'coloniality of data relations' poses. We must remember here that – as Indigenous and Black scholars have long told us – information and data have always been tools of the colonial project. Indeed, Ian Pool calls 'the collection, use and misuse of data on indigenous people' colonialism's (and postcolonialism's) 'fellow traveller' (2016, p. 57) – always, though, met by resilience and joyful resistance (Carlson and Frazer, 2021; Brock, 2020). It is very important to keep this history in mind when Big Tech appears to make similar moves on the historically privileged among us.

A potentially serious problem with some of these forms of critique is that they risk echoing, mirroring or amplifying – rather than debunking – the dominant myths about the power of technology. This risk is heightened when critique becomes unmoored from

specific, lived experience, and instead is discursively elevated to the same heights of global omniscience as Big Tech.

The clearest and most prominent example of Big Critique's tendency to mirror the rhetoric of Big Tech is Harvard business professor Shoshana Zuboff's epic – indeed, as anthropologist Anush Kapadia (2020, p. 33) has called it, 'operatic' – work *The Age of Surveillance Capitalism* (2019). In the book, Zuboff sounds the alarm about a new, aberrant form of capitalism that takes the excess data traces (or 'digital exhaust') left behind by our increasingly digitalised lives, converts them into predictive analytics, and then uses them for behavioural manipulation. Despite its overheated rhetoric, there is no doubt that works such as Zuboff's have highlighted increasingly urgent issues impacting society at both the individual and community levels and at planetary scale, and strong polemical language may aid in shifting the needle in more ethical directions (for an excellent 'review of the reviews' of *Surveillance Capitalism* that arrives at a similar conclusion, see Jansen and Pooley, 2021).

At the most polemical and populist end of the continuum, Big Critique is exemplified by the highly popular Netflix documentary film *The Social Dilemma*, released in mid-2020, and representing a particularly acute moment in a media environment awash in discourses of selfie narcissism, screen addiction, viral fake news and algorithmic radicalisation. The film centres on social media's algorithmic manipulation of user behaviour, and centres agency and responsibility almost entirely with the tech companies that provide some of the most popular social media platforms, primarily relying on tech insiders to diagnose the problems they claim (in hindsight) to have caused.

Despite the benefits of gathering attention for the cause, arguing from within the dominant framing (a pathological one, of addiction and behavioural manipulation), as both *Surveillance Capitalism* and *The Social Dilemma* do, risks becoming politically unproductive, because it discursively strips internet users – and all 'ordinary people' – of human agency. Not coincidentally, this framing also betrays the ongoing encroachment on media discourse of information-centred and behaviourist models of communication, so that we humans are cast as either the sinister agents or the unconscious subjects of behavioural tracking, targeting and manipulation – and that, of course, is exactly how Zuboff's 'surveillance capitalism' wants to see us. Therefore, in meeting hype with counter-hype (or, as Science, Technology and Society scholar Lee Vinsel [2021] calls it, 'criti-hype'), Big Critique can end up effectively reinforcing the claims that Big Tech makes about itself.

The pattern of media and industry hype and corresponding counter-hype around new technologies has a long history. The idea of a (permanent) technology 'revolution' is in turn articulated to deeply colonial ideas about the relationship between nationhood, progress and technology – particularly in the context of the United States, but with local resonances elsewhere. This cultural formation, which attributes awesome (or fearsome) power to AI, is the latest iteration of what David Nye (1996) called the American 'technological sublime', a framework later applied to the 1990s Silicon Bubble by Vincent Mosco (2005). Under this framework, the dominant narratives around data and AI as being all-powerful form part of the quasi-religious formation that is American technoculture (Nye, 1996; Carey and Quirk, 1970; Mosco, 2005; Brock, 2020), which represents itself as transcending lived experience,