

Secret agents: Algorithmic culture, Goodreads and datafication of the contemporary book world

European Journal of Cultural Studies

1–20

© The Author(s) 2019

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/1367549419886026

journals.sagepub.com/home/ecs

Simone Murray 
Monash University, Australia

Abstract

Goodreads, the Anglophone world's dominant book-centric social networking platform, is a compelling example of algorithmic selection of cultural goods. By exploring in detail Goodreads's corporate history, financing arrangements and commodification of user data, the article poses questions about the designed opacity of algorithmic selection processes, their self-perpetuating cultural effects, and potential privileging of the commercial interests of corporate owner Amazon. More broadly, the article ponders the optimal theoretical and methodological tools for examining the 21st-century book world. It ponders the shortcomings of standard book history approaches and canvasses what cultural and media studies frameworks may add. Given the increasing interpenetration of bookish dispositions and digital technologies, the article argues it is time for these disciplines themselves to merge.

Keywords

Algorithmic culture, Amazon, big data, black box, book history, cultural and media studies, datamining, Goodreads, publishing, recommendation

That we are now turning to algorithms to identify what we need to know is as momentous as having relied on credentialed experts, the scientific method, common sense, or the word of God.

– Tarleton Gillespie, 'The Relevance of Algorithms' (2014: 168)

Corresponding author:

Simone Murray, School of Languages, Literatures, Cultures and Linguistics, Monash University, Clayton, VIC 3800, Australia.

Email: Simone.Murray@monash.edu

The much-heralded death of the book at the hand of digital media has largely failed to eventuate. Feverish predictions during the 1990s that the codex book faced imminent eclipse by the rise of digital technologies have not come to pass (Bolter, 1991; Delaney and Landow, 1994; Nunberg, 1996; Spender, 1995). Granted, eBooks have made significant inroads in genre-writing niches, especially in the realm of online self-publishing, and all publishers have had their minds concentrated upon the pre-eminent importance of digital rights. Yet wholesale replacement of print has foundered on readers' residual fondness for the paper book, rival eBook retailers' strong-arm attempts to coerce customers into proprietary walled gardens, and concerns over an industry shift from outright ownership of a book copy to the eBook content-licencing model (Striphas, 2010a). While the statistical auguries can be difficult to interpret conclusively, eBook sales appear to have plateaued for some years now or are even in decline (Cain, 2017; Earls, 2017; Kottasová, 2017; *The Economist*, 2014). Yet any bibliophilic triumphalism at 'peak digital' having been reached would be sorely misplaced (Jenkins, 2016). While the central tenet of 1990s-style digital futurism – the death of the book – has failed to materialise, digital processes and platforms undeniably infiltrate the global book industry at every stage: from production (digital files, eBook rights, print-on-demand, online self-publishing, Wattpad, crowdfunded publishing), through circulation (online book retailing, authorial social media use, publisher search-engine optimisation, book trailers, blog tours, audiobooks), to consumption (reader reviews, fan fiction, bookish social networking, amateur booktubing, bookstagramming). The object of the printed book thus remains in wide circulation, but it is created, promoted, sold, evaluated, consecrated, consumed and debated within a pervasively digital agential mesh. Analysts of the contemporary book world thus need to cease conceptualising the analogue and digital as ontological opposites and instead examine the two domains' complex patterns of coexistence, mutual dependence and even, counterintuitively, revivification. Most recently, the rise of algorithmically-powered discoverability has again recast inherited processes of the book trade, reconfiguring the relationship of its stakeholders to each other and to newer entrants from the realm of Silicon Valley.

The pervasive influence of algorithms on contemporary society is well documented and constitutes a rich and expanding topic of academic research (Andrejevic et al., 2015; Beer, 2009; Eubanks, 2017; Gillespie, 2014; Kitchin and Dodge, 2011). Justifiably, a prominent concern has been automated news content and the democratic implications of a generalised public sphere splintering into evermore fractious and self-affirming filter bubbles (Bucher, 2018; Just and Latzer, 2017). More generally, scholars have investigated the role of algorithms in constructing the *cultural* horizon of 21st-century Internet users, with studies of audio-visual content selection (in particular YouTube and Netflix) and recorded music (e.g. The Echo Nest, Spotify) proliferating (Hallinan and Striphas, 2016; Morris, 2015; Pasquale, 2015). Yet the potential significance of algorithmic culture specifically for the world of books and reading has seldom been broached (cf. Chalmers and Edwards, 2017; Davis, 2015; Lang, 2012). This might be because the self-identified bookish could appear among the least digitally native population subgroup – more Luddites than early adopters. Yet this is clearly a misconception, not only because of the pervasive digitisation of the entire book industry but also because the book world was, after all, the locus of Amazon's first foray into online retailing – the industry-wide

cataloguing device of the International Standard Book Number (ISBN) providing a handy extant tool for streamlining customer ordering, retail fulfilment and inventory management (Striphas, 2009: 102). In fact, the book world's history of cataloguing meta-data and the data-ready nature of printed content have long proven lures for new industry entrants, from Enlightenment publishers' codification of knowledge via the encyclopaedia, through the devising of 'scientific' library classification systems and mainstreaming of bestseller rankings, to technology companies such as Google's controversial book digitisation schemes.

In bringing together studies of the contemporary book world and digital media studies – two disciplines with scant academic dialogue¹ – this article asks how the discipline of print culture studies can borrow productively from media and cultural studies' analyses of big data to understand the altered dynamics of literary culture in contemporary digital environments. Conversely, how might the digitally-enabled literary community provide media studies with a particularly resonant example of long-standing medium loyalty and deep emotional engagement with cultural properties influencing online behaviours among a minority of generally well-educated, affluent and unusually articulate users? The present discussion takes Goodreads, the globally dominant book-centric social-cataloguing site (owned by Amazon since 2013), as its case study. It takes a political economy approach to unpacking Goodreads – its origins, revenue structure, and legal frameworks – as a powerful though under-examined player in the contemporary book industries. The website's beguiling abundance of actual reader responses to books has obscured for scholars the limited extent to which users either understand or can influence its algorithmic operations, leading to overblown claims of readerly empowerment. Compelling evidence of reading's contemporary resilience and freely available research archive though it may be, Goodreads is above all else a node in platform capitalism. Against this background, the article pursues two, inter-related lines of argument: analytically it probes the cultural implications of algorithmic selection for bookish diversity, and the status of readers as unwaged content-creators; theoretically it poses larger questions about how datafication challenges book history's underpinning conceptual models and methodological protocols. What is the significance of Amazon's various algorithms as invisible intermediaries in the contemporary book realm? How are analysts of print culture to account for non-human decision-making processes that are at once so manifestly powerful in market terms yet so veiled from scholarly scrutiny?

Algorithmic culture

While the study of software has been foundational to the discipline of computer science since its inception, it has fallen to media studies to move discussion from a predominantly applied focus on optimising computer programming for processing speed, scalability and user-friendliness to specifically *critical* investigation of how ubiquitous software shapes societies and social actors in particular ways and with differential effects.² It has, therefore, chiefly been scholars from the loose assemblage of media/communication/cultural studies who have examined the cultural, social, political and economic effects of software, specifically as they affect relations between users on one hand and corporate and/or state entities on the other, and also between users themselves.

Invoking its dual humanities and social science lineages, media/cultural studies has pioneered critical, sociologically embedded research into information and communication technologies. In book historian and media studies scholar Matthew Kirschenbaum's (2009) neat encapsulation, 'software studies and its kin are the collision of computer science and cultural studies'.

Roughly five years into social media's new-millennial hegemony, scholars were registering concern that the generalised euphoria around Web 2.0's increased 'democratisation' of content creation and 'empowerment' of everyday users functioned to obscure new, insidious power relationships deriving from algorithms' deep enmeshment in daily life – the 'complex underweave of power at play in the digital mundane' in David Beer's (2009: 999) phrasing. Beer invokes sociologist Scott Lash's work on 'post-hegemonic power' (p. 991) to sketch a paradigm for analysing the 'technological challenges to human agency offered by the decision-making powers of established and emergent software algorithms' (p. 986). It was an overdue call for media studies to account for the power of algorithms in constituting users' lived reality and to become explicitly political in its analyses of this new AI-generated paradigm. Beer's sense of the urgency of these questions encompasses both the harvesting of individuals' data and the potentially self-reinforcing effects of automated collaborative filtering on users' cultural selection: 'it is also important that we consider how the information provided by users, and other "similar" users, might affect the things they come across' (p. 997). Atypically for a media studies researcher, Beer explicitly mentions the book world in a passing reference to automated Amazon recommendations, positing these as an example of the calculated algorithmic filtering of commodities and information to which users are seemingly serendipitously exposed (p. 997; see also Morris, 2015: 448).

Most recently and pertinently for the current discussion, bookish communication scholar Ted Striphas has explored the specifically cultural implications of algorithmic selection in a series of articles and blogposts. His focus is on how the 'sorting, classifying, and hierarchizing [of] cultural artifacts' by algorithms increasingly pervade 21st-century social and cultural life, from Google search results through Netflix recommendations to Facebook newsfeeds (Hallinan and Striphas, 2016; Striphas, 2010b, 2012, 2015). It is the outsourcing of long-standing human processes of cultural filtering to an algorithm whose selection criteria are unavailable for questioning due to the 'black box' effects of proprietary software, trade secrecy, non-disclosure agreements and commercial-in-confidence embargoes that gives scholars in this area most cause for concern (Andrejevic et al., 2015: 379–380; Cheney-Lippold, 2011: 176; Neff et al., 2012: 304; Pasquale, 2015; Striphas, 2015: 406–407). Moreover, because such algorithms operate in a digital environment of manifest cultural plenitude, their rhetorical positioning is always as a helpful guide through over-abundance – allaying consumer 'choice paralysis' by promising to individually customise cultural selection. 'I objectively and with impeccable neutrality merely serve up the items similar customers have most enjoyed', the algorithm seems meekly to promise, all the while effacing its own role in selecting and thereby perpetuating the display of selected items.

There is a very real risk of overstating algorithms' independence from human will, exacerbated by such ventriloquising of the algorithm as I have just indulged in. Algorithms may 'survey, capture, and process information about people and things in automated,

automatic, and autonomous ways, making judgments and enacting outcomes algorithmically without human oversight' (Kitchin and Dodge, 2011: x). Yet such algorithms are manifestly still coded by human developers, tested by IT programmers on actual user data, and constantly monitored and fine-tuned by human operators (Gillespie, 2016; Morris, 2015).³ Even the data an algorithm classifies and sorts must itself be made ready for capture by human choices. But at least during the period when an algorithm first 'goes live', it functions automatically and without direct human intervention in its decision-making – the fabled source of algorithms' speed, processing clout and accuracy. Artificial intelligence is at play here, one not entirely subject to human will. It is not necessary, in order to apprehend the significance of algorithms, to accord them the status of full intentionality and absolute autonomy. As social theorist Bruno Latour (2005) remarks, that would merely amount to 'a rather silly argument about the casual agency of technical objects', veering perilously close to technological determinism (p. 70). It is sufficient that algorithms are, for practical purposes, 'functionally automatic' (Gillespie, 2014: 170).

In his 2009 article, Beer sketches a useful tripartite schema for future Web 2.0 research which is intriguing in closely mirroring analogue media studies' long-standing division into familiar production/text/consumption frames of analysis (p. 998).⁴ Somewhat paraphrased here, Beer sketches a future research trajectory capable of examining:

1. The political economy of Web 2.0 companies (ownership, financing, data-harvesting practices, third-party licencing);
2. The functioning of algorithmic software (user data collected, how it is classified, metadata's influence on search results); and
3. User behaviours (knowledge of and responses to the algorithm, attempts to consciously cultivate online self-image, strategies to game the system).

Such broad-scale understanding of the task confronting cultural and media studies researchers is especially pertinent as it intersects with ongoing research into the contemporary book world. Scant prior academic work on Goodreads (e.g. Driscoll and Rehberg Sedo, 2019; Matthews, 2016) has been almost exclusively concerned with the third level of user behaviours, effectively overestimating user power by overlooking the structural and programming concerns of the first two levels. While the paucity of publicly available information about the working of Goodreads' algorithm may account for bypassing of the second level, there is no justification for overlooking the corporate origins, aims, operations and sale of Goodreads, given the amount of information about the firm on the public record – albeit usually couched in the habitually celebratory discourse of business journalism. The benefit of models such as Beer's is to refocus concern on how the ubiquity of algorithms in daily life throughout the developed world leads consumers actively or passively to consent to surveillance of their activities as well as harvesting and classifying of the data trails they leave (Striphas, 2010a). The corporate motivation for such extensive and intensive data gathering is to classify users into highly niche tribes for marketing and advertising purposes – whether resulting data are used by the Web 2.0 firm itself or licenced to third parties (Kitchin and Dodge, 2011: 184). It is crucial that analysts of the contemporary book fully apprehend the impact of such processes, as the

world of books has frequently cast itself as refuge from the perils and pervasive commercialisation of digital media. Though this pose is manifestly false given the intensive digitisation of all aspects of the book industries, the fact that readers frequently still choose to consume books in the comfortably analogue codex format allows the pretence to continue. More than this, big data companies predicated on the datafication of readers' bibliophilic selves and their social networks are only too happy to pander to the idea of literary culture as hallowed respite from cynically surveilling social media – thus disingenuously having their cake and eating it too.

Book history goes digital

None of the foregoing should suggest that researchers of the contemporary book world have been somehow oblivious to the medium's increasing digital enmeshment over the quarter century since the mainstreaming of the Internet. The study of the book not for its content (long corralled as literary studies' domain) but for its format and the intellectual, social, political and commercial implications of that format have long been the concern of the discipline known as book history. Borrowing from early communication and media studies work by Harold Innis, Marshall McLuhan and Walter Ong, and freely cross-blending this with insights deriving variously from bibliography, cultural history and literary studies, book history has coalesced as a discipline with its own international scholarly association – the Society for the History of Authorship, Reading and Publishing (SHARP) – a suite of publication outlets including the annual journal *Book History*, and a long-running annual conference programme. In a sure sign of its achieved disciplinary identity and institutional visibility, book history boasts an edited *Reader* of essential texts (now in its second edition) and numerous curriculum-defining introductory textbooks.⁵ Each of these includes, usually in a concluding section or chapter, some discussion of the implications of digitisation for the book.

The problems with book history in its current state as a disciplinary base for conceptualising the datafication of the contemporary book world are fourfold. First, the discipline remains, despite the efforts of a ginger group within SHARP, overwhelmingly historical in its academic purview, generally preferring to understand the digital as the scanning, compilation and curation of print-born artefacts for research convenience.⁶ Second, its ancestral tributaries of literary studies, cultural history and physical bibliography have put it at institutional and intellectual remove from newer humanities groupings such as media and cultural studies, towards which book historians still evince unjustified wariness.⁷ Third, and most apposite for present purposes, since book history's emergence as an academic discipline in the early 1980s, it has internalised human-centred models of the book's lifecycle. In Robert Darnton's (1990 [1982]) classic 'Communications Circuit', a node-and-link model which undergirds book history theoretically, the various agents of Author, Publisher, Printers, Shippers, Booksellers and Readers are assumed to be rational, self-maximising individuals who collectively gate-keep a linearly conceived book culture.⁸ Thomas R. Adams and Nicolas Barker (1993) later proposed an influential revision of Darnton's circuit, countering the historian Darnton's prioritising of individuals with a bibliographer's sense of the primacy of the book object. Nevertheless, their various 'events' in the life of a book – 'publishing,

manufacturing, distribution, reception and survival' – similarly assume human agents driving these abstract processes (p. 15). So too do 21st-century re-imaginings of Darnton's circuit such as Padmini Ray Murray and Claire Squires' (2013) 'Digital Publishing Communications Circuit', which rightly notes the elision and blurring of various of Darnton's gatekeeper roles as a result of disintermediation, such as book retailer Amazon also acting as publisher and host of customer reviews on its website. Ray Murray and Squires further incorporate new bookish intermediaries into their model, such as literary agents, wholesalers, distributors, and eReader manufacturers in overdue recognition that these also profoundly influence the shape of contemporary book culture (pp. 5, 6, 8). Nevertheless, their model remains premised on the idea of the book trade as the aggregate functioning of embodied, human intelligence, and one moreover open to scrutiny by book historians.

By contrast, in algorithmic culture, haunted as it is by the designed opacity of the 'black box' effect, scholars can only speculate about the workings of proprietary algorithms such as Amazon's collaborative filtering software, at best attempting to reverse-engineer their logics by drawing inferences from various inputs and outputs (Morris, 2015: 457, 459). This inscrutable character of algorithms presents book historians with significant practical problems of transparency in attempting to account for software's impact on contemporary book culture. Clearly the leviathan that is Amazon exerts immense influence on the global book trade, but how are scholars to document, much less critique, algorithmic culture's self-reinforcing effects on cultural selection if denied access to the workings of the algorithm's engine-room? This methodological dilemma foregrounds an assumption until now latent in book history – namely, researchers' Panglossian expectation that the artefactual records of gatekeeping processes will, given the passing of sufficient time, be made available for scholarly scrutiny. The book historian's accustomed forms of evidence – archives, correspondence, diaries and manifestos – are typically under greater threat from being simply discarded by publishing companies preoccupied with their current financial situation or being cherry-picked for correspondence from famed authors for sale at auction (and, even then, frequently to research institutions which might be expected to make them available to scholars in turn). By contrast, for Big Data corporations, data have never been conceived of as simply a by-product of day-to-day business operations; it is the firm's prime asset, its corporate *raison d'être*. Proprietary data are thus unlikely ever to be regarded as sufficiently historical to be discarded or gifted to public institutions. It follows that book history cannot simply transpose its human-centred schemas and public-interest assumptions to a digital environment premised on a data-hoarding commercial paradigm. The discipline must theoretically reposition artificial intelligence as itself agentic in the Latourian sense of '*any thing that . . . modif[ies] a state of affairs by making a difference*' (Latour, 2005: 71; italics original).⁹ More specifically, book historians must add software recommendation systems to their long-standing complement of Bourdieusean cultural intermediaries – 'infomediaries', as cultural studies scholar Jeremy Wade Morris (2015) terms them – which crucially intervene between author and reader through their digital curation of book culture (p. 452).

A fourth and final caveat for book historians acclimatising themselves to the digital domain: because of the discipline's typically retrospective gaze, book historians have

likely unconsciously internalised a comfortable sense of dissociation from the matters they study. Of course, the socio-political and ideological debates in which books are both actors and of which they remain material embodiments continue to have contemporary ramifications (that much cultural historian Darnton bequeathed the field).¹⁰ But there is little sense that in undertaking the process of research book historians are impacting that which they study. Compare this to the digital environment, in which every click by a researcher on Amazon's database affects the display of items, not only in that user's customised webpage ('reihung' for each individual on the basis of purchasing and search history) but cumulatively for all other users also (Cheney-Lippold, 2011: 165). There can be no safe quarantining of researcher from object of analysis when algorithmic processes track our every digital encounter and reformulate displayed results accordingly (Marres, 2017: 185). As in the Schrödinger's cat thought experiment beloved of theoretical physicists, merely by observing the digital book world we are ineluctably influencing it.

Goodreads: algorithmic culture at work in the contemporary book world

Goodreads is the world's dominant book-centric social networking and self-cataloguing platform. Based in San Francisco, the company was launched in January 2007 by software developer Otis Chandler and Elizabeth Khuri (who later married). Chandler, who remains the firm's CEO, had previously built a social networking and dating site, and he recounts his conversion to literary match-making as a eureka moment (Juergen, 2011; Narula, 2014):

One afternoon while I was scanning a friend's bookshelf for ideas, it struck me: when I want to know what books to read, I'd rather turn to a friend than any random person or bestseller list. So I decided to build a website – a place where I could see my friends' bookshelves and learn about what they thought of all their books.¹¹

Originally one of a number of reading-centric Web 2.0 sites including Library Thing and Shelfari (both now partially or entirely Amazon-owned), Goodreads grew steadily from its inception with membership ballooning from 7 to 17 million users during 2012 (Fidelman, 2012; Narula, 2014).¹² Not coincidentally, in March 2013 Chandler announced that Goodreads was 'joining the Amazon family' for a rumoured US\$150 million (Spillman, 2013), although Amazon's ownership is nowhere explicitly flagged on Goodreads' website. Goodreads now indisputably dominates online literary sociability with over 65 million users, 2 billion books catalogued, and 68 million reviews.¹³ Not without justification, Goodreads proclaims itself 'the world's largest community of readers'.¹⁴

Goodreads' appeal to users is essentially threefold. First, it facilitates literary self-cataloguing via personalised 'bookshelves' of books read, currently being read, to be read, or all-time favourites. The allure of self-quantification, of curating one's reading life for a global audience of bibliophiles, proved central to Goodreads' success (Pandell, 2016). Second, the site hosts book reviews and ratings (both personal and cumulative), facilitating readers' transition from passive consumer to amateur critic. Third, Goodreads

provides individually customised book recommendations generated via collaborative filtering of other readers' preferences (Thelwall and Kousha, 2017: 974).¹⁵ Never short of a punchy pitch-line, Goodreads styles itself 'the Netflix of book recommendations'.¹⁶ Tellingly, this functionality is everywhere presented as a flatteringly personalised service to the searcher, not as the harvesting of the searcher's data for other readers' convenience nor, more fundamentally, in the commercial interests of the company itself.

Delving beneath Goodreads' rhetoric of book-loving community reveals a business model based primarily on dual revenue streams of advertising and data licencing. For example, web analytics firm Quantcast's statistics for Goodreads reveal astonishingly detailed information about site users extending well beyond frequency of visits to the site.¹⁷ Data are compiled and displayed at granular level about users' geographical location (down to specific city), income bracket, educational attainments, ethnicity, household ownership, fashion preferences, media interests, other favourite websites, even their most likely brand of car to drive. From this, we glean a composite 'algorithmic identity' of the typical Goodreads user: a 25- to 34-year-old, US-based, Caucasian, graduate-educated woman with children, a median income of US\$100,000–150,000, who watches MSNBC and PBS, and enjoys science programmes (Cheney-Lippold, 2011: 165). So far, so predictable, perhaps. But more tellingly, Quantcast psychometrics micro-categorise Goodreads visitors into precisely targetable tribes according to their extra-literary consumer preferences such as a tendency to be early adopters of new technology, likely mortgage refinancers, or rusted-on credit-card loyalists. For Goodreads, books serve as a lure of affiliation for a particular well-educated, culturally invested, and affluent demographic of a kind immensely attractive to advertisers.¹⁸ By this logic, Goodreads' prime customers are in fact its advertisers (Nakamura, 2013: 241), with Goodreads members *themselves* constituting the site's key product.

Yet viewed in the context of Goodreads' other main revenue stream, the Review Partner Programme, users morph from commodity to free labour pool. Goodreads licences member-created book reviews to third parties including publishers, newspapers and Google Books for display on their websites.¹⁹ In media interviews, Chandler repeatedly extols Goodreads reader reviews as 'hands down the most amazing thing of Goodreads' [*sic*], verifiable evidence of the bookish community's gift economy (Narula, 2014). Yet, in Goodreads's pitch to would-be data-licensors, the emphasis is on the utility of amateur review content to lend an air of authenticity to others' websites, and thus to drive user stickiness and, ultimately, purchases:

Nothing sells a book like a thoughtful review written by a real person. Adding Goodreads reviews to your ecommerce site gives your customers more information, allowing for deeper interaction and a more profound connection with your brand. That creates a more satisfying customer experience – and more sales for you!²⁰

Despite 'members' being endlessly exhorted to maintain and curate their virtual bookshelves and reviews, the site's Terms of Use spell out unambiguously that users have no right to the data or profile they create on the website. Goodreads may cancel a user's account at any time 'for any reason or no reason' and without any liability, compensation or refund.²¹ Such asymmetry is typical of non-negotiable social media Terms of Use:

users are solely liable for any loss or damage arising from use of the site but simultaneously surrender all ownership and financial claim on the data they laboriously create or inadvertently leave (Nakamura, 2013: 241–242). Manifest here is Tiziana Terranova's (2000) definition of 'free labor', albeit now rendered in literary guise: 'productive activities that are pleurably embraced and at the same time often shamelessly exploited' by digital corporations which reap market value and brand circulation from the affective labour of the voluntariat (pp. 33, 37; also Kuehn and Corrigan, 2013: 10–11; Striplas, 2010a: 304).²²

Taken together, Goodreads users' demographic profiles and their emotional investment highlight precisely why Amazon was eager to purchase this rampant book industry rival. Chandler was already on record boasting 'we want to own book discovery' (Fidelman, 2012). By 2012, Goodreads was eclipsing Amazon on the measure of book browsers successfully converted to purchasers, with Goodreads's 29 percent conversion rate dwarfing Amazon's mere 10 percent. Chief among Goodreads's assets was its mainline into the reading habits of highly active book readers. Figures from the time reveal that a subsection of 19 percent of US adults are responsible for 79 percent of books read annually (Weissmann, 2013). Goodreads's near-monopoly on this group of opinion-influencing early adopters could thus be leveraged to promote mainstream book-purchasing patterns among less active readers (Vinjamuri, 2013a, 2013b). Despite early protestations to the contrary, within 3 years of the acquisition Amazon was combing Goodreads users' 'to be read' bookshelves to selectively promote discount Amazon eBooks, as formalised with the 2016 launch of Goodreads Deals (Pandell, 2016). Such strategic monetization of Goodreads data by 'the everything store' Amazon had been forecast at the time of the sale's announcement. Rob Spillman (2013), editor of (now defunct) literary magazine *Tin House*, confessed in forehead-slapping self-accusatory mode: 'Did it never occur to us Goodreads members that what seemed like a book-lover's paradise was actually a fantastically valuable chunk of pure data just ripe for the mining?'.

Attending to real readers

Beer's tripartite schema for future work on algorithmic culture takes care to balance the deterministic tendencies of political economy with cultural studies' openness to the potential subversiveness of actual user behaviours. After all, too much talk of the unsailable power of algorithms can have the counterproductive effect of enhancing algorithmic mystique. While some academic work exists on users' motivations and practices in engaging with the Goodreads website, such research has frequently misperceived the exact nature of Goodreads review content, and the power relationships that permeate it. Online reading scholar David Dowling's (2014) description of 'digital reading communities' as examples of 'deinstitutionalized participatory culture' strikes an especially disconcerting note. While reading on Goodreads floats free of geographically based or formal educational institutions, readers' very contributions to online reading discussions – their posts, reviews and social networks – are *themselves* the stuff of Goodreads's institutional value. Granted, the Internet has indisputably 'unleashed the culture from the institutions that formerly regulated intellectual exchange'. But that does not prevent born-digital corporations converting newly digitally hosted intellectual exchanges into

readily and exclusively commercialisable data (Dowling, 2014). In so doing, Goodreads and other Amazon subsidiaries firmly entrench themselves as the dominant institutions of 21st-century literary culture (McGurl, 2016).

Similarly, researcher Jolie C. Matthews (2016) declines to explore commodity relations between Goodreads and its users' data, only interactions between groups of users on the website. She acknowledges that amateur reviews 'have become a commodity' in the sense that they are liked or circulated by other users (p. 2314). Furthermore, honest and fair reviewing indisputably aids in the social evaluation of material commodities by helping readers to decide whether a particular book is worth their limited time and money. But by focusing on the metaphorical commodity value of Goodreads reviews as reputational markers or purchaser-advice, Matthews overlooks how Goodreads' Terms of Use 'agreement' *literally* commodifies users' reviews and other browser data under highly restrictive intellectual property (IP) arrangements.²³ The net effect is not, *pace* Dowling, *de*-institutionalisation so much as a changing of the book world's institutional guard.

Writing more recently, Beth Driscoll and DeNel Rehberg Sedo (2019) (from literary and media studies backgrounds, respectively) acknowledge the commercial utility of Goodreads reviews and note that prior work about the platform has frequently focused on such production dimensions (pp. 248–249). In seeking instead to read Goodreads reviews themselves in aggregate and note performances of various kinds of intimacy, they deploy feminist standpoint theory to dignify the self-understanding of readers. However, the sociological and cultural studies disciplines from which such approaches arise tend to conceive of female emotion, especially in its 'excessive' forms, as somehow subversive of a privileged order – even if only semiotically so. On Goodreads, highly emotional discourse, even when it is explicitly conflictual (as where two reviewers heatedly and protractedly disagree about a title) is commercially beneficial in that it encourages users to prolong their involvement with the site. Rather than constituting a form of readerly resistance of the kind long celebrated in both book history and cultural studies, displays of emotional intimacy on Goodreads are commercially highly serviceable. If, as Driscoll and Rehberg Sedo concede, the commercial function of reviews to the publishing industry 'is not evident in most Goodreads reviews', it is harder still to frame online readerly behaviours as self-conscious subversion of their host site (p. 257). Such methodologies have yet to make clear how intense readerly sentiment differs from ideal consumer behaviour and, if it does not, what its scholarly utility might be.

Interpreting the 'social reading graph'

Outlining Goodreads's corporate growth and its ever-closer integration into Amazon's operations illuminates the site's functioning in broader industry and cultural spheres.²⁴ Goodreads as case study thus serves a dual function here: specifically, it showcases the impact of Web 2.0 on the contemporary book world, as well as emblematising microcosmically the datafication of cultural selection processes generally. Automated cultural selection depends upon algorithms that are extremely costly to create, refine and maintain, and which thus constitute closely guarded corporate IP (Gillespie, 2014; Just and Latzer, 2017; Striplhas, 2010b). Goodreads in fact found developing an effective

book-recommendation algorithm in house such a ‘really, really hard problem to get right’ that in March 2011 it purchased another company, Discovereads, purely for its algorithms (Halford, 2011).²⁵ Because such proprietary algorithms constitute ‘big data’ firms’ chief asset, the workings of such algorithms are deliberately kept opaque. The market value of an algorithm (and, by extension, its parent company) depends upon its reputational validity. Logically, there is a greater risk of a website being ‘gamed’ if details of the algorithm’s workings are made public, so companies have every incentive to enhance the mystique of ‘the Algorithm’ (in truth typically a collection of multiple, interdependent algorithms) by withholding design and operational specifics.

One salient feature of algorithms as examples of AI is that they do not merely present snapshots of data at a given point in time, but are constantly learning – using information acquired from past processing to refine and guide future decision-making (Bucher, 2018: 14). Every time we rate our books read, log our to-be-read lists, or even click on other titles, we provide the algorithm with additional data to fine-tune its portrait of our bibliophilic selves (Gillespie, 2014: 173). A key effect of machine-learning’s use of past results to condition future results is a high degree of self-perpetuation, a dynamic especially pronounced and problematic in programmes facilitating specifically *cultural* selection. Algorithmically generated reading suggestions resemble past books read, only more so. Algorithms thus create normative models of readerly consumption – ‘statistical stereotypes’ as it were – to which readers are encouraged to conform ever more closely (Cheney-Lippold, 2011: 171). In Galloway’s (2004) phrasing, algorithms ‘determine and at the same time inflect the identity of the user’ (p. 114). This is strikingly at odds with the conventional representation of literary history as a series of aesthetic revolts against ossified tradition, even with the received timeline of literary criticism as the throwing over of tired intellectual schemas by bold Young Turks.

Even at the individual level, algorithms’ remorselessly self-perpetuating logic goes against the grain of readerly desire: readers seek not only confirmation of existing tastes but challenges to and reorienting of those tastes (Cheney-Lippold, 2011: 169). Hence, merely reinforcing a reader’s existing taste profile is self-defeating in that it ‘engineers spontaneity out of the picture’, in the words of James Marcus (2004) – memoirist, early Amazon Literary Editor, and self-styled humanist crusader against digital Fordism inside the Seattle-based empire (p. 157, 200). As author and *New Yorker* reviewer Macy Halford (2011) has more recently observed about Amazon’s now subsidiary, Goodreads:

I could go on clicking to the end of time or a hundred and ninety million titles, whichever comes first, rating not just books I’ve read but those I ‘want’ to read. The algorithm would improve with each click and together we’d progress toward the completion of the game, a moment when we’d both win: a library precisely tailored to my desires would burst forth, an information ecstasy. But that point is vanishingly far away. By the time I’d got there, my reading preferences would have changed and I’d have to start over again.

The technologically self-confirming tendencies of AI have a propensity to be reinforced by aggregate human decision-making. As demonstrated by quantifying tools such as Google’s PageRank (ranking most-visited and linked websites) and BookScan (logging retail book sales), software which merely claims to count cultural products or information

tends to be readily adopted as a de facto arbiter of cultural worth (Haigh, 2006; Knox, 2005). In a book industry long starved of accurate data – where publishing decisions were frequently guided by the ‘print-and-pray’ variety of market research – quantitative metrics have been belatedly seized upon as *qualitative* measures of value (Rachman, 2018). What appears at the top of such listings is most frequently emulated and therefore stays at the top of the listing, effectively polarising a cultural field into a handful of mega ‘hits’ and a very long tail of also-rans. In a further self-perpetuating loop, cultural producers or consumers latch onto quantitative data which merely purports to document current trends and use these predictively to guide future choices such as book commissioning, title acquisition, retail stock orders, and to-be-read lists (Davis, 2015: 515). The net effect is to filter cultural offerings, tending to aesthetic conformity, a narrowing of the marketplace of ideas, and the contraction of an individual’s horizon of possibility.

A third and final caveat for book world scholars in engaging with the ubiquity of algorithmic culture is the potential for the commercial interests of an algorithm’s owners to inflect results. Corporate owners of algorithms do not offer quantifying services benevolently; they are in the business of selling products and services (principally consumer goods in Amazon’s case, advertising in the case of Facebook). Especially because of the ‘black box’ effect, there is no way of knowing whether an algorithm is skewed to favour in-house offerings, nor what impact decisions by the corporate owner may have on the hierarchy of results displayed. Tarleton Gillespie (2014) notes that ‘algorithms [*are*] perennially open to user suspicion that their criteria skew to the provider’s commercial or political benefit’ yet, because of designed opacity, this suspicion can never be definitively confirmed, either by everyday users or specialist researchers (p. 176). In a book industry with a narrow and shrinking range of retail outlets this concern is especially pronounced because of the long-running eclipse of independent bookstores by chains, the more recent collapse of some of those chains, and Amazon’s hegemony over online retailing. Amazon’s blatantly punitive behaviour towards business ‘partners’ that cause it commercial displeasure, such as the repeated removal of the ‘buy’ button from titles by publishers with which it was in dispute, hardly dispels suspicions of self-interested software tinkering (Robinson, 2010; Stone and Rich, 2010). On-the-record precedents like these seem to confirm Frank Pasquale’s (2015) judgement that Internet and finance firms ‘cloak self-serving appraisals and conflicts of interest in a veil of technological wizardry’ (p. 9). Importantly from the perspective of academic researchers, cultural algorithms do not altruistically present us with objective, arms-length market data. Despite their scrupulously maintained appearance of pure functionality and blameless neutrality, algorithms in truth permit only contingent access to in-house intelligence in the interests of building market share for the company and socialising users into a habit of daily reliance on the software (Gillespie, 2016). Free access can be, and in the case of corporate Facebook accounts has been, revoked unilaterally by corporate fiat and with scarcely any notice.

Conclusion: ‘ripe for the mining’?

What insights does detailed examination of Goodreads’s readerly datafication yield for contemporary book studies, cultural/media studies and the possible inter-relationship of

the two disciplines? For some decades now it has been orthodox in studies of the contemporary book world to invoke Pierre Bourdieu's model of the cultural 'field' (*champ*) for its ability to abstractly represent the dynamic interplay of multiple agents and their often competing agenda. However, Bourdieu's very glancing attention to the Internet, even in his late works, makes it a matter of urgency for book studies to expand its intellectual range to incorporate media and cultural studies research into algorithmic culture. While addressing a pressing theoretical problem, such a turn towards datafication raises a new, specifically methodological dilemma: How to account for the power of algorithms when those algorithms are unavailable for scholarly scrutiny, likely in perpetuity? To ascertain a way out of this dilemma it is worth reverting to the original 'black box' model. While the contents of the black box can only be speculated about, inferences can be drawn from a system's verifiable inputs and outputs (Andrejevic et al., 2015: 380). This evokes levels 1 and 3 of Beer's model for research into Web 2.0, making it essential to scrutinise the corporate structures, revenue sources and business partnerships of algorithmic cultural decision-makers on one hand, as well as users' perceptions, strategic behaviours and cultural commentary (such as they are) on the other. To some extent, researchers have begun embarking on the latter quest, but the former remains book studies' veritable Dark Continent.

Undeniably, the majority of users continue to engage with Goodreads because of the utility and pleasure the site affords them. But as professional analysts of the cultural and media industries, it is incumbent upon us to critique Goodreads's self-interested implicit positing of the book as somehow outside of and in opposition to digital culture, when the site's own Terms of Use belie such a confection. In this, the very bookishness of sites such as Goodreads gives analysts a head start. For the book world has long understood itself to be a site of resistance to the complete capture of culture by commerce: most of the convulsing book-world conflicts of recent decades hinge upon the lingering cultural associations of books that somehow exempt them from reduction to mere consumer commodities. On the other hand, the very existence of book history is premised on dispelling literary studies' fiction that texts are merely abstract verbal collations and instead insisting on critically examining written communication's embodied and material substrates. The specific challenge for book history is that the datafication of the contemporary book world is at once profoundly *de*-materialised in that its forms evade bibliography's habitual minute physical categorisation of the codex format. Yet at the same time, it is ineluctably materialised in the sense of enmeshing books in a world of vast profits, corporate IP and data-mining. Doing justice to the book's newly datafied existence will require scholars with a foot in both (rapidly converging) book history and cultural/media studies camps.

Funding

The author received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Simone Murray  <https://orcid.org/0000-0001-6213-2146>

Notes

1. Ted Striphas is one of few scholars working at the intersection of these two disciplines who is affiliated with a university Communication programme and publishes primarily in media, communication and cultural studies journals (refer, for example, Striphas, 2010a). Otherwise, scholars working in this liminal space frequently identify primarily as book historians or literary studies specialists (refer Note 6).
2. See Alexander R. Galloway (2006).
3. On the issue of algorithms' relative independence from humans and the possibility of 'technical agency' see, respectively, Striphas (2015: 408, n.1) and Neff et al. (2012: 305).
4. In similar vein, Gina Neff et al. (2012), writing some years later, urged scholars to consider 'how [software] tools are designed, how they function socially, and how users are aware of their positions and power' (p. 301).
5. Refer David Finkelstein and Alistair McCleery (2006); David Finkelstein and Alistair McCleery (2013); Leslie Howsam (2006); A. R. Hawkins (2006); Solveig C. Robinson (2014); Michelle Levy and Tom Mole (2017) and James Raven (2017).
6. Overwhelmingly does not, of course, mean entirely. Books drawing on both book history and contemporary cultural and media studies frameworks include Matthew G. Kirschenbaum's (2008) *Mechanisms*, Anouk Lang's (2012) edited anthology *From Codex to Hypertext: Reading at the Turn of the Twenty-First Century*, and my own *The Digital Literary Sphere: Reading, Writing, and Selling Books in the Digital Era* (Murray, 2018). Scholars publishing article- or chapter-length discussions in this interdisciplinary zone include Daniel Allington; Mark Davis; Beth Driscoll; Mark McGurl; DeNel Rehberg Sedo; Simon Rowberry; Claire Squires; and Ann Steiner. Also analysing the contemporary book world more from a cultural sociology perspective are scholars John B. Thompson, Laura J. Miller and C. Clayton Childress. Such broadly conceived 'publishing studies' approaches still constitute a minority at SHARP conferences, though a notably growing one.
7. Refer Leslie Howsam's (2006) useful triangular diagram of book history's intellectual ancestry (p. 10).
8. Since first publication in journal form in 1982, the article has garnered 1014 citations according to Google Scholar (current at time of writing). Darnton is also book history's top-cited scholar by a wide margin.
9. In related vein, Leah Henrickson (2018) has recently considered the implications of attributing authorial agency to natural language generating (NLG) computer systems producing literary works.
10. Note that talismanic book history pronouncements such as SHARP co-founder Jonathan Rose's (2003) epigrammatic 'Books make history, and history makes books' already attribute agency to the book object (p. 11). Given this, it is even harder to rationalise the resistance to extending such agency (if not intentionality) to computer algorithms.
11. <https://www.goodreads.com/about/us>
12. Amazon acquired a 40 percent stake in Library Thing in 2006 and purchased Shelfari outright in 2008. In 2016, Amazon folded Shelfari into Goodreads (to the consternation of some vocal long-term Shelfari users).
13. <https://www.goodreads.com/about/us>
14. <https://www.goodreads.com/jobs>
15. <https://www.goodreads.com/>
16. <https://www.goodreads.com/blog/show/303-announcing-goodreads-personalised-recommendations>
17. <https://www.quantcast.com/goodreads.com>. Since research for this article first began in February 2018, Quantcast has restricted access to its metrics to paying subscribers.

18. <https://www.goodreads.com/advertisers>
19. https://www.goodreads.com/about/review_program
20. https://www.goodreads.com/about/review_program. Commentators such as Yesha Naik (2012) note the pervasiveness of Goodreads reviews elsewhere online, but fail to attribute this to their on-selling by Goodreads for profit (p. 321).
21. <https://www.goodreads.com/about/terms>
22. While Terranova (2000) was writing years before the advent of Goodreads, she gives as examples of ‘immaterial labor’ the work of ‘writing/reading/managing and participating in mailing lists/Web sites/chatlines’ (p. 42). Furthermore, consider this as a prescient thumbnail sketch of Goodreads’ rationale and current functionality: ‘Users keep a site alive through their labor, the cumulative hours of accessing the site (thus generating advertising), writing messages [reviews], participating in conversations [book clubs], and sometimes making the jump to collaborators [handpicked “First Reads” reviewers]’ (p. 49). Goodreads users have been alerted to their wilful self-exploitation by pseudonymous Zoe Desh (2016), severely disgruntled creator of the *Goodreads Sucks* website and its self-published companion booklet *Authors vs Goodreads: A Cautionary Tale*: ‘Members of Goodreads are fools to provide this lucrative subsidiary of Amazon with untold thousands of hours of free labor trying to maintain the unmanageable mess their much hyped *billion+book* database is in’ (<http://goodreads-sucks.com/>).
23. The implications of highly restrictive Amazon and Goodreads Terms of Use agreements are explored in more detail in chapter 4 of my *The Digital Literary Sphere: Reading, Writing, and Selling Books in the Internet Era* (Murray, 2018).
24. <https://www.goodreads.com/api>
25. <https://www.goodreads.com/blog/show/271-recommendations-and-discovering-good-reads>

References

- Adams TR and Barker N (1993) A new model for the study of the book. In: Barker N (ed.) *A Potencie of Life: Books in Society*. London: British Library, pp.5–43.
- Andrejevic M, Hearn A and Kennedy H (2015) Cultural studies of data mining: Introduction. *European Journal of Cultural Studies* 18(4–5): 379–394.
- Beer D (2009) Power through the algorithm? Participatory web cultures and the technological unconscious. *New Media and Society* 11(6): 985–1002.
- Bolter JD (1991) *Writing Spaces: The Computer, Hypertext, and the History of Writing*. Hove; London: Lawrence Erlbaum.
- Bucher T (2018) *If. . . Then: Algorithmic Power and Politics*. New York: Oxford University Press.
- Cain S (2017) Ebook sales continue to fall as younger generations drive appetite for print. *The Guardian*, 14 March. Available at: <https://www.theguardian.com/books/2017/mar/14/ebook-sales-continue-to-fall-nielsen-survey-uk-book-sales> (accessed 15 August 2018).
- Chalmers MK and Edwards PN (2017) Producing ‘One Vast Index’: Google Book search as an algorithmic system. *Big Data and Society* 17(2): 1–16.
- Cheney-Lippold J (2011) A new algorithmic identity: Soft biopolitics and the modulation of control. *Theory, Culture & Society* 28(6): 164–181.
- Darnton R (1990 [1982]) What is the history of books? In: Darnton R (ed.) *The Kiss of Lamourette*. London: Faber, pp.107–135.
- Davis M (2015) E-books in the global information economy. *European Journal of Cultural Studies* 18(4–5): 514–529.
- Delaney P and Landow G (1994) *Hypermedia and Literary Studies*. Boston, MA: The MIT Press.
- Desh Z (2016) *Authors vs. Goodreads: A Cautionary Tale*. Los Gatos, CA: Smashwords.

- Dowling D (2014) Escaping the shallows: Deep reading's revival in the digital age. *Digital Humanities Quarterly* 8(2). Available at: <http://www.digitalhumanities.org/dhq/vol/8/2/000180/000180.html> (accessed 6 March 2018).
- Driscoll B and Rehberg Sedo D (2019) Faraway, So Close: Seeing the Intimacy in Goodreads Reviews. *Qualitative Inquiry* 25(3): 248–259.
- Earls N (2017) Has the print book trumped digital? Beware of Glib conclusions. *ABC News*, 11 May. Available at: <http://www.abc.net.au/news/2017-05-09/has-the-print-book-trumped-digital-beware-glib-conclusions/8510984> (accessed 15 August 2018).
- Eubanks V (2017) *Automating Inequality: How High-Tech Tools Profile, Police and Punish the Poor*. New York: St. Martin's Press.
- Fidelman M (2012) These are top 25 book reviewers on Goodreads. *Forbes*, 16 October. Available at: <https://www.forbes.com/sites/markfidelman/2012/10/16/goodreads-ceo-these-top-25-book-reviewers-represent-the-future-infographic/#5029c5865fff> (accessed 23 March 2018).
- Finkelstein D and McCleery A (eds) (2006) *The Book History Reader* (2nd edn). London: Routledge.
- Finkelstein D and McCleery A (eds) (2013) *An Introduction to Book History* (2nd edn). New York: Routledge.
- Galloway AR (2004) *Protocol: How Control Exists after Decentralization*. Cambridge, MA: The MIT Press.
- Galloway AR (2006) *Gaming: Essays on Algorithmic Culture* (Electronic Mediations Series). Minneapolis, MN: University of Minnesota Press.
- Gillespie T (2014) The relevance of algorithms. In: Gillespie T, Oczkowski PJ and Foot KA (eds) *Media Technologies: Essays on Communication, Materiality, and Society*. Cambridge, MA: The MIT Press, pp.167–194.
- Gillespie T (2016) Algorithm. In: Peters B (ed.) *Digital Keywords: A Vocabulary of Information Society and Culture*. Princeton, NJ: Princeton University Press, pp.18–30.
- Haigh G (2006) Information idol. *The Monthly*, February, 25–33. Available at: <https://www.themonthly.com.au/monthly-essays-gideon-haigh-information-idol-how-google-making-us-stupid-170>
- Halford M (2011) Getting good at Goodreads. *The New Yorker*, 18 September. Available at: <https://www.newyorker.com/books/page-turner/getting-good-at-goodreads> (accessed 23 February 2018).
- Hallinan B and Striphas T (2016) Recommended for you: The Netflix prize and the production of algorithmic culture. *New Media and Society* 18(1): 117–137.
- Hawkins AR (ed.) (2006) *Teaching Bibliography, Textual Criticism, and Book History*. London: Pickering & Chatto.
- Henrickson L (2018) Tool vs. agent: Attributing agency to natural language generation systems. *Digital Creativity* 29: 182–190.
- Howsam L (2006) *Old Books and New Histories: An Orientation to Studies in Book and Print Culture* (Studies in Book and Print Culture). Toronto, ON, Canada: University of Toronto Press.
- Jenkins S (2016) Books are back. Only the technodazzled thought they would go away. *The Guardian*, 13 May. Available at: <https://www.theguardian.com/commentisfree/2016/may/13/books-ebook-publishers-paper> (accessed 15 August 2018).
- Juergen M (2011) How Goodreads got started. *Entrepreneur*, 22 February. Available at: <https://www.entrepreneur.com/article/218120> (accessed 23 March 2018).
- Just N and Latzer M (2017) Governance by algorithms: Reality construction by algorithmic selection on the Internet. *Media, Culture and Society* 39(2): 238–258.

- Kirschenbaum MG (2008) *Mechanisms: New Media and the Forensic Imagination*. Cambridge, MA: The MIT Press.
- Kirschenbaum MG (2009) Where computer science and cultural studies collide. *The Chronicle of Higher Education*, 23 January. Available at: <http://www.chronicle.com/article/where-computer-science-and/14806> (accessed 15 September 2017).
- Kitchin R and Dodge M (2011) *Code/Space: Software and Everyday Life* (Software Studies Series). Cambridge, MA: The MIT Press.
- Knox M (2005) The ex factor: BookScan and the death of the Australian novelist. *The Monthly*, May, 51–55. Available at: <https://www.themonthly.com.au/books-malcolm-knox-exfactor-bookscan-and-death-australian-novelist-34>
- Kottasová I (2017) Real books are back. E-book sales plunge nearly 20%. *CNN*, 27 April. Available at: <https://money.cnn.com/2017/04/27/media/ebooks-sales-real-books/index.html> (accessed 15 August 2018).
- Kuehn K and Corrigan TF (2013) Hope labor: The role of employment prospects in online social production. *The Political Economy of Communication* 1(1): 9–25. Available at: <http://www.polecom.org/index.php/polecom/article/view/9>
- Lang A (ed.) (2012) *From Codex to Hypertext: Reading at the Turn of the Twenty-First Century*. Amherst, MA: University of Massachusetts Press.
- Latour B (2005) *Reassembling the Social: An Introduction to Actor-Network-Theory*. Oxford: Oxford University Press.
- Levy M and Mole T (2017) *The Broadview Introduction to Book History*. Guelph, ON, Canada: Broadview Press.
- McGurl M (2016) Everything and less: Fiction in the age of Amazon. *Modern Language Quarterly* 77(3): 447–471.
- Marcus J (2004) *Amazonia*. New York: New Press.
- Marres N (2017) *Digital Sociology: The Reinvention of Social Research*. Cambridge: Polity.
- Matthews JC (2016) Professionals and nonprofessionals on Goodreads: Behavior standards for authors, reviewers, and readers. *New Media and Society* 18(10): 2305–2322.
- Morris JW (2015) Curation by code: Infomediaries and the data mining of taste. *European Journal of Cultural Studies* 18(4–5): 446–463.
- Murray S (2018) *The Digital Literary Sphere: Reading, Writing, and Selling Books in the Internet Era*. Baltimore, MD: Johns Hopkins University Press.
- Naik Y (2012) Finding good reads on Goodreads: Readers take RA into their own hands. *Reference and User Services Quarterly* 51(4): 319–323.
- Nakamura L (2013) ‘Words with Friends’: Socially networked reading on Goodreads. *PMLA* 128(1): 238–243.
- Narula SK (2014) Millions of people reading alone, together: The rise of Goodreads. *The Atlantic*, 12 February. Available at: <https://www.theatlantic.com/entertainment/archive/2014/02/millions-of-people-reading-alone-together-the-rise-of-goodreads/283662/>
- Neff G, Jordan T, McVeigh-Schultz J, et al. (2012) Affordances, technical agency, and the politics of technologies of cultural production. *Journal of Broadcasting and Electronic Media* 56(2): 299–313.
- Nunberg G (ed.) (1996) *The Future of the Book*. Berkeley, CA: University of California Press.
- Pandell L (2016) Goodreads is finally cashing in on its devoted community. *Wired*, 19 May. Available at: <https://www.wired.com/2016/05/goodreads-selling-books/> (accessed 16 February 2018).
- Pasquale F (2015) *The Black Box Society: The Secret Algorithms That Control Money and Information*. Cambridge, MA: Harvard University Press.

- Rachman T (2018) The agony and the ecstasy of taking author photos. *Literary Hub*, 19 March. Available at: <https://lithub.com/the-agony-and-the-ecstasy-of-taking-author-photos/> (accessed 20 March 2018).
- Raven J (2017) *What is the History of the Book?* Cambridge: Polity.
- Ray Murray P and Squires C (2013) The digital publishing communications circuit. *Book 2.0* 3(1): 3–23.
- Robinson C (2010) The trouble with Amazon. *The Nation*, 14 July. Available at: <http://www.thenation.com/article/37484/trouble-amazon> (accessed 18 July 2014).
- Robinson SC (2014) *The Book in Society: An Introduction to Print Culture*. Peterborough, ON, Canada: Broadview Press.
- Rose J (2003) The horizon of a new discipline: Inventing book studies. *Publishing Research Quarterly* 19(1): 11–19.
- Spender D (1995) *Nattering on the Net: Women, Power and Cyberspace*. Melbourne, VIC, Australia: Spinifex Press.
- Spillman R (2013) Amazon buys Goodreads: We're all just data now. *Salon*, 1 April. Available at: https://www.salon.com/2013/03/31/amazon_buys_goodreads_were_all_just_data_now/ (accessed 16 February 2018).
- Stone B and Rich M (2010) Amazon removes Macmillan books. *The New York Times*, 30 January. Available at: <http://www.nytimes.com/2010/01/30/technology/30amazon.html> (accessed 18 July 2014).
- Striphas T (2009) *The Late Age of Print: Everyday Book Culture from Consumerism to Control*. New York: Columbia University Press.
- Striphas T (2010a) The abuses of literacy: Amazon Kindle and the right to read. *Communication and Critical/Cultural Studies* 7(3): 297–317.
- Striphas T (2010b) How to have culture in an algorithmic age: Algorithmic culture, electronic reading. *The Late Age of Print Blog*, 14 June. Available at: <http://www.thelateageofprint.org/2010/06/14/how-to-have-culture-in-an-algorithmic-age/> (accessed 23 March 2018).
- Striphas T (2012) What is an algorithm? *Culture Digitally*, 1 February. Available at: <http://culture-digitally.org/2012/02/what-is-an-algorithm/> (accessed 23 March 2018).
- Striphas T (2015) Algorithmic culture. *European Journal of Cultural Studies* 18(4–5): 395–412.
- Terranova T (2000) Free labor: Producing culture for the digital economy. *Social Text* 18(2): 33–58.
- The Economist (2014) The future of the book: From papyrus to pixels. *The Economist*, 11 October, 45–50.
- Thelwall M and Kousha K (2017) Goodreads: A social network site for book readers. *Journal of the Association for Information Science and Technology* 68(4): 972–983.
- Vinjamuri D (2013a) The trouble with finding books online – And a few solutions. *Forbes*, 27 February. Available at: <https://www.forbes.com/sites/davidvinjamuri/2013/02/27/the-trouble-with-finding-books-online-and-a-few-solutions/#78fb23d92720> (accessed 23 February 2018).
- Vinjamuri D (2013b) Three hidden benefits of the Amazon acquisition of Goodreads. *Forbes*, 29 March. Available at: <https://www.forbes.com/sites/davidvinjamuri/2013/03/29/three-hidden-benefits-of-the-amazon-acquisition-of-goodreads/#4f3995c731f5> (accessed 23 February 2018).
- Weissmann J (2013) The simple reason why Goodreads is so valuable to Amazon. *The Atlantic*, 1 April. Available at: <https://www.theatlantic.com/business/archive/2013/04/the-simple-reason-why-goodreads-is-so-valuable-to-amazon/274548/> (accessed 16 February 2018).

Biographical note

Simone Murray is associate professor in Literary Studies at Monash University, Melbourne where her research centres upon sociologies of literature. Her book *Mixed Media: Feminist Presses and Publishing Politics* (Pluto Press UK, 2004) was awarded the 2005 SHARP DeLong Book Prize. Her second monograph, *The Adaptation Industry: The Cultural Economy of Contemporary Literary Adaptation* (Routledge, 2012), has been widely reviewed in English-, French-, German- and Swedish-language publications. Her latest monograph, *The Digital Literary Sphere: Reading, Writing, and Selling Books in the Internet Era* (Johns Hopkins UP, 2018), examines how the Internet has transformed literary culture.