
PASTEUR4OA Briefing Paper: Infrastructures for Open Scholarly Communication

Authors: Samuel Moore (OKF, King's College London and Ubiquity Press), Jonathan Gray (OKF) and Danny Lämmerhirt (OKF)

Reviewers: Alma Swan (EOS)

June 2016



Introduction

Over the past two decades, open access advocates have made significant gains in securing public access to the formal outputs of scholarly communication (e.g. peer reviewed journal articles) – through a combination of public policy-making, funder mandates and changing the norms and practices of researchers in different fields. The same period has seen the rise of platforms from commercial publishers and technology companies that enable users to interact and share their work, as well as providing analytics and services around scholarly communication. How should researchers and policy-makers respond to the rise of these platforms? Do commercial platforms necessarily work the interests of the scholarly community? How and to what extent do these proprietary platforms pose a threat to open scholarly communication? What might public alternatives look like? This paper provides a brief overview of the rise of scholarly platforms – describing some of their main characteristics as well as debates and controversies surrounding them. It argues that in order to prevent new forms of enclosure, it is essential that public policy-makers should be concerned with the provision of public infrastructures for scholarly communication as well as public access to the outputs of research. It concludes with a review of some of the core elements of such infrastructures, as well as recommendations for further work in this area.

Platformisation of the Web

In 2004 Tim O'Reilly described the evolution of Web 2.0 as 'the web as platform' (O'Reilly 2004). The web was no longer a place for simply publishing/consuming information; it was now promoted as a collaborative, dynamic and 'participatory' web (Beer 2009). In this vein, the web became a place for user-generated content and sites such as Facebook, YouTube and Twitter allowed their users to upload, share and interact with each other's content – creating an audience of so-called 'prosumers', simultaneously becoming producers and consumers. Platforms have since evolved into centralised locations for a range of user interactions and the web is now dominated by a handful of large companies that have each platformised their own area of the market: e.g., Google for search, Amazon for retail, Facebook for social networking, and even Uber for transportation. Platforms tend to be:

1. **Data-generating:** The currency of this new realm is data. Platforms often base their business models around monetising transactional data between users, Platforms are designed to both increase the size of its user-base and maximise the number of interactions in the given network,

and the design of the platform itself strongly influences the kinds of interactions that take place. As studies on social networking platforms show, the platform design (like buttons, re-tweet function) serves to capture user-generated data that are sold to clients such as marketing departments. (See Bechmann (2013) for more on the data-generating strategies of Google and Facebook.)

2. **Multi-stakeholder-management:** Platforms cater different services to different user groups, often serving to cross-subsidise free access business models. Platforms such as Facebook and Twitter offer free use of their platform to anyone. The platform design (like buttons, re-tweet function) serve to capture user-generated data that are paid content for clients such as marketing agencies. The cross-subsidisation offering different services to different clients is a common trait of “free” platforms. But critics may argue that this causes asymmetries in access to data that would be relevant for the larger public – such as public universities interested in researching these user interactions.
3. **Profit-maximising:** Platforms are almost exclusively for-profit entities, often funded by venture-capital funding. For this reason, many do not establish business models until *after* they have acquired a substantial user-base, thus making them a more attractive acquisition for larger companies who seek to incorporate users onto their own platforms. An example is Twitter which, despite having hundreds of millions of users and billions in venture capital funding, has still not reached profitability.
4. **Walled Gardens:** In seeking to maximise the number of users and interactions on their site, platforms can be best described as walled gardens. They seek to keep the user within its own central ecosystem to the greatest extent possible, as interactions occurring outside the platform cannot be monetised. For this reason platforms are also monopoly-seeking and try to extend their services to as many domains as possible so that users need never leave the platform’s ecosystem (e.g., Google’s single sign-in for its vast array of products). Platforms may use APIs to permit third-party app development via an application programming interface or API (Helmond 2015). This allows developers to access the platform’s data and features to create new applications that in turn drive users to the host platform. While APIs do represent gates into these walled gardens, they are still routinely subject to the centralised governance of the platform owners, rather than the open, decentralised web.

The dominance of platforms represents a significant shift in the architecture of the web to a more proprietary ecosystem of a small number of large players, each designed to attract and retain users, to capture and quantify usage, and to cater services for free to one client group, while offering additional paid services to other client groups. This shift has considerable implications for the infrastructures that support open access (OA) to research outputs, as will now be described, and a move away from platforms and towards public infrastructures will be recommended.

Open Access and Platformisation

In parallel to the platformisation of the web, the past decade also saw the widespread uptake of open access to scholarly research. Much has been written on the move to OA (and the PASTEUR4OA resources provide a detailed overview of the various topics¹) though the chief focus of the literature on OA is on the strategies and business models for providing access to peer-reviewed books and journals.

However, scholarly communication systems include much more than just formal publishing channels. More informal channels like email lists, blogs, grey literature, data repositories and social media all play an important role in the creation, dissemination and preservation of research. These informal communication channels are an important element in scholarly communication, including the pre-publication of grey literature or other papers (a function that is offered by Social Science Research

¹ <http://www.pasteur4oa.eu/resources> [accessed 18/06/16]

Network). Importantly, communication infrastructures reflect certain ‘politics’ – they can be designed in accordance with the values of open access, allowing users to self-archive, access and discuss content with other researchers in the network (as SSRN does). These politics are different for commercial scholarly platforms, posing questions if these infrastructures are able to support open access communication.

As scholarly communications become more open, but the web becomes more populated with walled gardens, it is wise to interrogate some of the current platforms that govern and shape these more informal channels for OA – this section outlines the features of a few key scholarly platforms:

Google Scholar

Google Scholar is a widely adopted service for research discovery and tracking citations, pulling content from ‘trusted sources’, including repositories, institutional pages, journals, open-access indexes, etc., and suggesting related results where relevant. Drawing on Google’s powerful search functionality, Google Scholar is well populated with research from all disciplines, much of it being freely accessible to the user. Search results are returned in order of how useful Google believes the content is, relative to the search terms:

Google Scholar aims to rank documents the way researchers do, weighing the full text of each document, where it was published, who it was written by, as well as how often and how recently it has been cited in other scholarly literature.²

However, Google does not provide any specific information on how these criteria are judged and how results are presented to the user. Max Kemman describes two of these effects as disproportionately promoting older papers and non-elite journals (Kemman 2015). Researchers are thus beholden to Google’s closed algorithm, a hallmark of centrally governed platforms as described above, and there are no clear processes for dispute resolution, nor any clear incentives for Google to provide such a service. This is also important in the context of user profiles, which display a user’s publication list, co-authors and citations accrued, and Google offers minimal direct assistance for users hoping to fix a problem (and instead asks users to contact the publisher of the article in question³).

Academia.edu/ResearchGate

There are a number of social networking sites for academics, most notably ResearchGate⁴ and Academia.edu⁵. These sites offer many of the familiar features of social media, such as personal profiles, connecting with other users, and analytics about profile and document views. Each site is built around users sharing their publications, allowing them to seek feedback on drafts and share published papers – the platforms therefore act as both repositories and discovery services, but with a social layer for facilitating interactions. Both platforms are well adopted (ResearchGate claims 10 million users while Academia.edu boasts almost 40 million) and their features are clearly valued by researchers. However, like many platforms, the lack of sustainable business models associated with such services means their future direction is often unclear.

What is clear is that these free services will be valued on the sizes of their user-bases and the extent to which they can, as Gary Hall argues, ‘exploit the data flows generated by the academics who use the platform as an intermediary for sharing and discovering research’ (Hall 2015). Both companies have

² <https://scholar.google.co.uk/intl/en/scholar/about.html> [accessed 18/06/16]

³ <https://scholar.google.co.uk/intl/en/scholar/citations.html#citations> [accessed 18/06/16]

⁴ <https://www.researchgate.net/about> [accessed 18/06/16]

⁵ <https://www.academia.edu/about> [accessed 18/06/16]

received many millions in venture capital⁶ and will need to please their investors either through sustainable generation of profit or an exit (in the form of an acquisition). Much of ResearchGate's business strategy appears to be targeting advertising to its users, while Academia.edu's approach is less clear, leading some to suspect they are positioning to be bought. As Kathleen Fitzpatrick claims:

[A] limited number of options for the network's future: at some point, it will be required to turn a profit, or it will be sold for parts, or it will shut down. (Fitzpatrick 2015).

Like Google Scholar, the lack of community governance means there is no guarantee that either ResearchGate or Academia.edu will be maintained with the best interests of the research community in mind – or even maintained at all, in the event of an acquisition.

Mendeley

Mendeley is an interesting example of a platform that began life as a 'full-fledged' open access platform before being acquired in 2013 by the Dutch publisher Elsevier. Initially it was a reference manager and allowed users to share papers via their public profiles – it was well adopted and highly regarded in the open access community. The site was also a vocal proponent of open data until its acquisition and granted free access to its usage data through its API. Since the acquisition, as Cameron Neylon argues, the public sharing of articles is 'long gone'; instead, openly available articles have been replaced with direct links to publisher versions (often behind a paywall). What's more, the API was initially neglected after the acquisition but has since seen an influx of investment, though the outputs of which Neylon describes as indicating Elsevier's 'desire to maintain control over information' in the form of proprietary, restricted data (Neylon 2016). This is seen as connected to the Social Science Research Network (SSRN), an open-access repository for social science preprints and another of Elsevier's recent business acquisitions.

The acquisition of SSRN

Elsevier is contested by large sections of the research community and it has been a vocal opponent of open access.⁷ As the official press release⁸ states Elsevier plans integrating the repository into its wider list of products and to cover new market segments⁹. As described above, Elsevier's objective here is usage data – they have acquired a substantial amount of content and hope to monetise this by applying the services of other products owned by Elsevier. An excerpt of the official acquisition statements tells:

Elsevier is actively linking data and analytics to its vast content base in ways no other potential SSRN partner can match. By connecting Mendeley, Scopus, ScienceDirect and its editorial systems, it is helping researchers get a more complete picture of their research landscape. Institutions will also benefit with a better view of their researchers' impact.¹⁰

Use data and impact analytics are a commodity to generate profit from free scholarly platforms. Institutions which are increasingly in the need to quantify the performance of their researchers may

⁶ Information from: <https://www.crunchbase.com/> [accessed 18/06/16]

⁷ <http://sparcopen.org/news/2015/new-policy-from-elsevier-impedes-open-access-and-sharing/> ; <http://www.theatlantic.com/science/archive/2016/01/elsevier-academic-publishing-petition/427059/> ; <http://www.michaeleisen.org/blog/?p=807> ; <http://thecostofknowledge.com/> [all accessed 26/06/2016]

⁸ See the official announcement: <https://www.elsevier.com/connect/ssrn-the-leading-social-science-and-humanities-repository-and-online-community-joins-elsevier>

⁹ A blogpost comment argues that "because it is "open access" and because it collects an impressively diverse range of social science in one place, SSRN's data actually represents the world of social science scholarship reasonably well. See also: <http://savageminds.org/2016/05/18/its-the-data-stupid-what-elseviers-purchase-of-ssrn-also-means/>

¹⁰ <https://www.elsevier.com/connect/ssrn-the-leading-social-science-and-humanities-repository-and-online-community-joins-elsevier>

financially depend on scholarly platforms when buying access to the usage data and impact metrics of their scholars.

The above examples are illustrative of the kinds of well-utilised services for scholarly communication that have been shaped by platformisation. They exemplify many of the general characteristics of platforms described in the first section and are representative of the web's turn toward walled gardens, and the problems this entails. That said, the features of such services are clearly valued by the research community. Platforms have an incentive to build such features in order to attract more users, perhaps in a way that traditional businesses, non-profits and higher education do not. But there are greater concerns with how such for-profit, data-driven platforms will impact on the scholarly community more specifically, as the next section will describe. The paper concludes with a recommendation for new ways forward for research infrastructures and how the research community might facilitate this future.

Critiques and Concerns

Standards for Preservation and Open Access

Platforms often strategically promote 'openness' as their core feature – empowering their users for certain actions. While the concept may be a vague one, openness is an attractive quality that implies a kind of freedom and transparency.(see Tkacz 2014). Though Mendeley may have toned down its rhetoric surrounding open access, platforms generally and Academia.edu in particular all see themselves as promoting openness in many forms.¹¹ While this may be true (there is certainly more publicly accessible information available from platforms than subscription journals), there are certain standards associated with open access that are absent from these platforms. As is widely known, open access does not simply imply a research object that is free to access but one that is free to reuse too.¹² The ability to reuse an article tends to be granted by a Creative Commons licence, most commonly CC BY, which allows readers to share, translate and build upon scholarly works. In the absence of this licence, as is common on scholarly platforms, the kind of open access provided is severely lacking. This is exacerbated by the requirement of some platforms to sign-up for an account in order to download an article. Such features represent a serious dilution of the intentions of open access to provide freely and immediately provide the rights to read, use and re-use scholarly information.

The upshot is that most platforms should not be used as open-access repositories because of the control maintained by each platform. Katie Fortney and Justin Gonder provide an excellent summary of why academic social networks should not be seen as OA repositories (Fortney and Gonder 2015), from where the following graphic is taken (reproduced under CC BY):

¹¹ <https://googleblog.blogspot.co.uk/2009/12/meaning-of-open.html> ;
https://www.academia.edu/12297791/Open_Access_Meets_Discoverability_Citations_to_Articles_Posted_to_Academia.edu ;
<https://www.academia.edu/press> [accessed 27/06/16]

¹² <http://sparcopen.org/open-access/>

	Open access repositories	Academia.edu	ResearchGate
Supports export or harvesting	Yes	No	No
Long-term preservation	Yes	No	No
Business model	Nonprofit (usually)	Commercial. Sells job posting services, hopes to sell data	Commercial. Sells ads, job posting services
Sends you lots of emails (by default)	No	Yes	Yes
Wants your address book	No	Yes	Yes
Fulfills requirements of UC's OA policies	Yes	No	No

(cc) BY <http://creativecommons.org/licenses/by/4.0/> University of California OSC

Academic social networks clearly do not conform to the standards that have been carefully devised by the institutional repository community over many years. They lack the necessary metadata standards, long-term preservation plans, interoperability and the ability to export content in bulk. Martin Eve also makes recommendations for ways in which Academia.edu can better serve the open access community, from allowing bulk data downloads and becoming interoperable with other repositories, to fostering greater understanding of the legal and moral side of open access (Eve 2014). This however may conflict with current business models relying on walled gardens and the exploitation of user data within a proprietary ecosystem – catered to certain clients who are willing to pay a certain amount of money.

This speaks to a broader point on the design of platforms. As is well known, the sciences receive substantially more funding than arts and humanities disciplines, making them more attractive sources of income. This is already on display in the adoption article-processing charges open access publishing that overwhelmingly favour grant-funded research. Although many platforms are used by a range of scholars from a number of disciplines, there is likely to be a concentration of features and services on the disciplines with the most money or from those whose data are more valuable to R & D companies, meaning the arts, humanities and humanistic social sciences will be considered second-class disciplines. Platforms are therefore not discipline-neutral; they are designed to ‘follow the money’.

Quantification of Research Assessment

The popularity of using metrics as a proxy for research quality is another noteworthy feature of scholarly platforms. Academic career advancement has always been governed by extraneous measures of quality such as the Journal Impact Factor, H-Index and journal reputation. The adoption of data-driven platforms for scholarly communication means there are more sources of data by which research(ers) can be judged. Indeed, as James Wilsdon et al. report in the recent independent review of metrics in higher education, quantitative data is now used ‘far more widely as a management aid’ throughout higher education (Wilsdon 2016). Grant income is tracked and measured against benchmarks, citations are counted using commercial platforms and teaching is evaluated through surveys. So-called altmetrics aim to establish new impact metrics for research – particularly those based on social media shares, blogs, news stories, etc. – that are necessary for individual career advancement and the evaluation of research

impact more broadly. There is a definite move towards the quantification of academia and the data may be provided by the kinds of platforms described above. Impact measurement and quantitative assessment of impacts are intimately linked to career advancement and the way academia evaluates itself. Exploiting this field opens new dependencies of public higher education institutions.

While such data may be useful in certain contexts, the research community needs to be aware of who controls it and how it is being calculated and presented. If the data are being sourced via commercial platforms, not only does the research community not own the data, it also has no control over which statistics are promoted over others (as some transactional data may be more lucrative than others). As Bilder et al. argue:

The scholarly community does not own or control most of this information. For example, we could have built or taken on the infrastructure to collect bibliographic data and citations but that task was left to private enterprise. Similarly, today the metadata generated in scholarly online discussions are increasingly held by private enterprises. They do not answer to any community board. They have no obligations to continue to provide services at their current rates, particularly when that rate is zero. (Bilder, Lin, and Neylon 2015).

But many take a more pessimistic view of the kinds of data generated by platforms. Indeed, higher education in many countries is already beholden to performance reviews based on more and more sources of data. As Stefano Collini writes in the *London Review of Books*:

The true use-value of scholarly labour can seem to have been squeezed out; only the exchange-value of the commodities produced, as measured by the metrics, remains. (Collini 2013).

New scholarly platforms simply create more ways by which academics can be measured, which in turn creates a need to obtain a competitive advantage over other academics, especially in such a difficult job market for higher education. Researchers have been forced to become 'microentrepreneurs', as Gary Hall terms them, each needing to maintain a public presence as a way of advertising oneself and one's research. Academic platforms thus create the conditions of possibility for a range of scholarly activities and certain new ideas get normalised: that researchers need to think of themselves as brands, that the division between work and leisure time should be blurred and that one's academic worth can be measured by the number of connections one has, the papers one has published and the citations one has accrued.

What is perhaps most pernicious, as Hall observes, is that it is near impossible for researchers (or anyone for that matter) to opt out of these services and refuse to supply platforms with their data:

Nevertheless, refusing to be part of this move toward supplying ever more quantities of information, data, and work for free is not an option for most people. It is not something that can be opted out of or strategically abandoned and withdrawn from simply by declining to look for research on Google Scholar, removing cookies from your computer, [and] committing social networking suicide [...]. (Hall 2016, 35).

This is the direction the web is heading in higher education and beyond. The ubiquity of data-driven systems has irrevocably changed the web, including scholarly communications specifically, and will continue to influence the ways in which open access is adopted throughout the research community. But are there possible alternatives? What might public infrastructures for scholarly communication look like?

Public infrastructures for scholarly communication

Scholarly infrastructures should be just as important to the research community as the more formal routes to open access. While business models are transitory and evolving, infrastructures maintain a level of permanence that often goes unquestioned. Drawing on the philosophy of Martin Heidegger, Tony Ross-Hellauer shows that infrastructures are invisible in everyday experience, that is, until they stop functioning as expected, much like the pen that only enters the writer's consciousness when it stops working or runs out of ink. For Ross-Hellauer, this means that 'infrastructure can be so near or so ever-present and yet is rarely fully perceived and reflected upon [...]' (Ross-Hellauer 2016). Given this, critical evaluation of scholarly infrastructures should be encouraged and new systems funded and adopted where possible.

Bilder, Lin and Neylon propose four key aspects to infrastructures for scholarly communication (paraphrased below), many of which are absent from the platforms described above:

1. Governance: They should be 'stakeholder governed, transparent and non-discriminatory, understanding they do not have a right to exist beyond the service provided to the community.
2. Sustainability: Infrastructures should be made to generate surplus, have contingency plans and, crucially, base their revenue on services not data.
3. Insurance: The research community needs to know it is in control and can maintain the control in the event of anything unexpected. The authors recommend that services be based on open-source software, open and reusable data and patent non-assertion. This will allow the community to inherit aspects of the service that for whatever reason can no longer be maintained.
4. Implementation: The trickiest aspect of these principles is how they are implemented and what kind of organisations are most suited to them. The authors assert that they naturally lead to a 'board-governed, not-for-profit membership organisation' but other models should be explored too, including centralised vs. federated systems. (Bilder, Lin, and Neylon 2015).

Naturally, many institutions already conform to a number of these principles – the point is to recognise and promote their positive aspects. Bilder and colleagues suggest that ORCID is closest to meeting these criteria. The Open Researcher and Contributor ID (ORCID) seeks to create one identifier for one researcher, thus disambiguating researchers, allowing an easy, de-centralised sign-in mechanism and connecting researchers to their list of publications. The organisation is a non-profit and releases its software as open source; it is governed by a board of executives (comprising researchers, publishers and technologists) and steering groups from business, communications and technology sectors.

Similarly, the EU-funded OpenAIRE project aims to provide and interlink some of these scholarly infrastructures. It funds Zenodo, hosted by CERN, which is an open-source repository for data, preprints, posters, software, and so on.¹³ OpenAIRE also offers a decentralised search function that allows users to search for research outputs across a range of institutional, subject and international repositories.¹⁴ OpenAIRE is governed by representatives from each EU member state and an academic advisory board. Both ORCID and OpenAIRE can be used as blueprint organisation models for public scholarly infrastructures. Policymakers should facilitate and incentivise these organisations to work together, promoting an ecosystem of mutual reliance and shared legal and technical frameworks.

However, it is clear that many of the valued aspects of platforms arise from their slick design, novel features and their resulting network effects. In order to attract users, platforms attract millions of venture capital that they can risk on innovative services and features, in ways that centrally or commonly owned alternatives might not be able to compete with. Private enterprise of course has a part to play in scholarly infrastructures and it is also unrealistic to assume that all scholarly infrastructures can be both publicly

¹³ <https://zenodo.org/> [accessed 28/06/16]

¹⁴ <https://www.openaire.eu/search/find?keyword=> [accessed 28/06/2016]

and centrally governed. However, new models of funding such operations are encouraged beyond venture capital and business models based on private ownership of data.

For example, Mariana Mazzucato in her book *The Entrepreneurial State* encourages governments to take more risk and invest in early-stage companies through a variety of measures, but also to reap and re-distribute the rewards (Mazzucato 2014). This would allow the state to have a say in how platforms are developed, monetised and eventually sold, creating a more responsible basis for innovation. In a similar vein, Michel Bauwens and Kostakis advocate the 'partner state approach' for the promotion of 'sustainable models of entrepreneurship and participatory politics' (Kostakis and Bauwens 2014); and Terbor Scholz and others advocate 'platform cooperativism' as a way to challenge the dominance of data-driven corporate platforms, allowing common ownership and governance for all who participate in the platform (Scholz 2016). While there is not the space to go into detail here, platforms created by these mechanisms would be able to ensure that the generated data is publicly owned and the platforms sensitively designed. Further work is therefore needed to explore new economic models for scholarly infrastructures.

Concluding Note

It is clear that the research community needs to be wary of adopting scholarly platforms that seek to generate and privatise control of user data. As more of the literature becomes open access, there are greater opportunities for commercial platforms to assert their dominance with 'value added' services for discovery, networking and self-promotion, all of which resulting in increased quantification of researchers themselves. In opposition to this, the community needs to look towards and nurture research infrastructures that permit greater governance, sustainability, and that operate in service of research itself, while still offering the features that researchers require. Failure to do so may result in the ceding control of our research infrastructures to a handful of large, disinterested companies for whom the research community is not the customer, but the product itself.

References

- Bechmann, Anja. 2013. 'Internet Profiling: The Economy of Data Intraoperability on Facebook and Google'. *MedieKultur: Journal of Media and Communication Research* 29 (55): 19.
- Beer, David. 2009. 'Power through the Algorithm? Participatory Web Cultures and the Technological Unconscious'. *New Media & Society* 11 (6): 985–1002.
- Bilder, Geoffrey, Jennifer Lin, and Cameron Neylon. 2015. 'Principles for Open Scholarly Infrastructures'. *Science in the Open*. <http://cameronneylon.net/blog/principles-for-open-scholarly-infrastructures/>.
- Collini, Stefan. 2013. 'Sold Out'. *London Review of Books*, October 24.
- Eve, Martin. 2014. 'Things Academia.edu (@academia) Should Do (hint: Work with Green OA)'. *Www.martineve.com*. <https://www.martineve.com/2014/09/05/things-academia-edu-academia-should-do-hint-work-with-green-oa/>.
- Fitzpatrick, Kathleen. 2015. 'Academia, Not Edu'. *Planned Obsolescence*. <http://www.plannedobsolescence.net/academia-not-edu/http://www.plannedobsolescence.net/academia-not-edu/>.
- Fortney, Katie, and Justin Gonder. 2015. 'A Social Networking Site Is Not an Open Access Repository'. *Office of Scholarly Communication: University of California*. <http://osc.universityofcalifornia.edu/2015/12/a-social-networking-site-is-not-an-open-access-repository/>.
- Hall, Gary. 2015. 'What Does Academia_edu's Success Mean for Open Access? The Data-Driven World of Search Engines and Social Networking'. *LSE Impact Blog*.

- <http://blogs.lse.ac.uk/impactofsocialsciences/2015/10/22/does-academia-edu-mean-open-access-is-becoming-irrelevant/>.
- . 2016. *Pirate Philosophy: For a Digital Posthumanities*. Cambridge: MIT Press.
- Helmond, Anne. 2015. 'The Platformization of the Web: Making Web Data Platform Ready'. *Social Media+ Society* 1 (2): 2056305115603080.
- Kemman, Max. 2015. 'Standing on the Shoulders of the Google Giant: Sustainable Discovery and Google Scholar's Comprehensive Coverage.' *LSE Impact Blog*.
<http://blogs.lse.ac.uk/impactofsocialsciences/2015/11/19/standing-on-the-shoulders-of-the-google-giant/>.
- Kostakis, Vasilis, and Michel Bauwens. 2014. *Network Society and Future Scenarios for a Collaborative Economy*. New York: Palgrave Macmillan.
- Mazzucato, Mariana. 2014. *The Entrepreneurial State: Debunking Public vs. Private Sector Myths*. London: Anthem Press.
- Neylon, Cameron. 2016. 'Canaries in the Elsevier Mine: What to Watch for at SSRN'. *Science in the Open*. <http://cameronneylon.net/blog/canaries-in-the-elsevier-mine-what-to-watch-for-at-ssrn/>.
- O'Reilly, Tim. 2004. 'Opening welcome: The state of the internet industry'. presented at the Web 2.0 Conference, San Francisco, California. http://web-2con.com/presentations/web2con/intro_tim_john.ppt.
- Ross-Hellauer, Tony. 2016. 'Infrastructure Is Invisible / Infrastructure Is Law : OpenAIRE Blog'. *OpenAIRE Blog*. <https://blogs.openaire.eu/?p=979>.
- Scholz, Trebor. 2016. 'Platform Cooperativism: Challenging the Corporate Sharing Economy'. New York: Rosa Luxemburg Stiftung.
- Tkacz, Nathaniel. 2014. *Wikipedia and the politics of openness*. Chicago: University of Chicago Press.
- Wilsdon, James. 2016. *The Metric Tide: Independent Review of the Role of Metrics in Research Assessment and Management*. SAGE.
<http://www.hefce.ac.uk/pubs/rereports/Year/2015/metrictide/Title,104463,en.html>.