Chapter 7. Big data and new social relations in higher education: Academia.edu, Google Scholar and ResearchGate

Janja Komljenovic

Lancaster University, Lancaster, United Kingdom, j.komljenovic@lancaster.ac.uk

This document is the final author's version as it will appear in: Sellar, S., Gorur, R. and Steiner-Khamsi, G. (Eds.). *World Yearbook of Education 2019: Comparative Methodology in an Era of Big Data and Global Networks*. London: Routledge.

Introduction

Digital technologies have become part of our everyday life in many different forms. Particularly, social media enable us to connect and communicate in novel ways across institutional or national borders. The higher education sector and academics are no exception. Researchers extensively use social media at all points of the research lifecycle – from identifying research opportunities to disseminating results (Mas-Bleda, et al., 2014; Rowlands, et al.,, 2011).

UNESCO reports that there are 7.8 million researchers in the world (UNESCO, 2015), while the founders of academic social media platforms estimate that there are about 17 to 20 million academic platform users altogether, a number that includes graduate students and other potential users (Duffy & Pooley, 2017; Scott, 2017). Like everybody else, these academics use digital platforms, post content, comments and likes, become friends and network with others, upload their papers, look at advertisements, look for jobs, and so on. While doing all this, they produce 'big data' and leave their digital traces. Digital technologies have thus made it possible to collect, store, analyse and use unprecedented amounts of data.

Are the digital platforms and the data they produce just new technological devices like others in the past, or do they represent something inherently new? Does this data enable new ways of knowing the social world and, if so, how? What does all this data do? And finally, what kind of methodologies can we apply in education studies to make use of this data?

These are questions that I will investigate in this chapter in relation to academics and the knowledge they produce. Big data, platforms and digital technology, on the one hand, present new data sources for social research, while on the other hand they are objects of enquiry themselves. By using the cases of three digital platforms targeted at academics, I will first explore what kind of research possibilities the platforms and big data enable for us methodologically. More specifically, I will examine what we can learn about academia by analysing the data produced by academic users of digital platforms. Second, I will focus on digital platforms as objects of enquiry and explore what kind of social relations these platforms construct and what kind of social world they constitute.

Social media platforms and academia

About social media platforms

Social media are defined as "websites and applications that enable users to create and share content or to participate in social networking" (Oxford Dictionary, 2016). They are commonly associated with the 'Web 2.0', which refers to more participatory and collaborative use of the Internet and the production and modification of online content by users. Websites today increasingly incorporate user generated content and social networking features, which are key characteristics of Web 2.0 technologies. By so doing they are enabling users to collaborate and bring their social activities to the Internet. What was previously individual searching, reading or referencing, is now a social activity where people share readings or their comments about these readings on academic and other platforms (Gruzd, Staves,

& Wilk, 2012). When people use social media to create and share content, they also produce large amounts of data.

Some authors see these unprecedented amounts of data as new sources of information about social life that could potentially replace traditional methodologies like surveying, while others argue that there is nothing inherently new about big data (Marres, 2017). Either way, the novelty of social media and wider digital technology in comparison to data collection about individuals by traditional means is that the personal record of data is "much, much more exhaustive, its components are processed automatically, and it circulates with much greater ease across institutions and markets" (Fourcade & Healy, 2017, p.11). Moreover, the distinguishing feature of the digital technologies of today are "their extensive capabilities for monitoring, analysing and informing social life" (Marres, 2017, p.8). In other words, it is not only about using social media as individuals and producing big data, but also about manipulating data and offering it back to users to inform or alter the social world.

Academic digital platforms

Traditionally, scholars have relied on printed academic publications and the peer review system to disseminate their research and communicate with others in their scholarly community. Now, academic publishing is thoroughly digital and the Internet has changed how scholars disseminate their research and communicate (Liu, 2003). Academic Web presence can be classified as personal, institutional and social (Mas-Bleda et al., 2014). Personal and institutional websites are created by individual scholars or universities to offer academic information; while social presence refers to platforms with automatically generated publication provision (such as Google Scholar and Microsoft Academic Search) or self-entry of publications (such as Mendeley, LinkedIn, Academia.edu and ResearchGate) (Mas-Bleda et al., 2014).

In this chapter, I focus on the three digital platforms that are specifically targeting academics for the purpose of sharing and discussing their research outputs or communicating in relation to their

research, namely: Academia.edu. Google Scholar Citations and ResearchGate.

Academia.edu

Academia.edu is a social media platform targeting academics, which was established in 2008 by Richard Price, then a doctoral student of philosophy at Oxford University. It has 54 million academic members, who added 19 million papers to the platform as of August 2017 (Academia.edu, 2017a). It positions itself as a platform "for academics to share research papers" and its mission is "to accelerate the world's research" (Academia.edu, 2017a). Richard Price states on his personal webpage that the goal of Academia.edu is to "get every science PDF ever written on the Internet, accessible for free. We want to make science faster and more open" (Price, 2017). Academia.edu started as a free service and it monetised particular features in December 2016 (Bond, 2017) to create "Academia Premium". It has attracted $17.7 million in investments by Khosla ventures, True ventures, Spark ventures and Spark capital (Academia.edu, 2017a).

The platform brings together individual users and papers. Users create a profile, share basic information and upload documents. Academia.edu does not offer a space for discussion like ResearchGate, but it does have an option for "sessions", where one can share drafts of papers and ask for feedback. When they log in to the platform, users receive personalised recommendations for which papers to read, which sessions to attend for commenting on draft papers and which academics to follow. Finally, they are subjected to commercial advertising and job advertising (Academia.edu, 2017b). Most recently, Academia.edu offers a "recommendation" function for users who are approved by the Editor Program to recommend papers.

Academia.edu is notable for one more reason. Duffy and Pooley (2017) report on the importance of Academia.edu's Web address ending with ".edu", which gives impression of a scholarly institution with non-profit credibility, while in fact it is a private company. This was possible because the domain was registered in 1999, before the 2001 regulations restricted the ".edu" designation to accredited higher education institutions.

Google Scholar Citations

Google Scholar is a platform that enables searching scholarly literature across disciplines and sources. It is the widest academic search engine as it includes journal and conference papers, theses and dissertations, academic books, pre-prints, abstracts, technical reports, and other scholarly literature. It ranks the documents for its search engine based on where a particular document was published, who it was written by, and the information on citation in other scholarly literature (Google Scholar, 2017a).

Google Scholar automatically creates an individual profile for its users, which then offers personalised information, namely information about user's citations, creating "my library" updates on papers and authors that the user is interested in, alerts on papers or authors that the user would like, and metrics of larger academic fields and papers (Google Scholar, 2017b).

Google Scholar Citations (known also as Google Scholar Profiles) was opened to the general public in 2011 (Martin-Martin et. al., 2016). It provides information to users about who is citing their publications, information on citations over time and information on several citation metrics. Users can also make their profiles public, which would lead to the profile appearing when people search for the user's name. Google automatically pulls information on papers and citations and allocates them to the user. The user can choose to have the updates performed automatically, check the proposed updates or do it manually (Google Scholar, 2017a).

ResearchGate

ResearchGate is a social media platform targeting researchers, which was established in 2008 by physicians Dr. Ijad Madisch and Dr. Sören Hofmayer, and computer scientist Horst Fickenscher. It has more than 13 million members as of August 2017 (ResearchGate, 2017a), which is 60 per cent of such potential users around the world based on Madisch's estimations (Scott, 2017). As of 2014, 14 million papers were accessible on the platform (Van Noorden, 2014) and it is reported that researchers upload 2.5 million papers to ResearchGate every month (Scott, 2017). Its mission is "to connect the world of

science and make research open to all" (ResearchGate, 2017a). It is free to its users and is financed through advertising for scientific products and services (Shead, 2017). It has attracted large investments from Benchmark Capital, Bill Gates, Tenaya Capital, Goldman Sachs, Wellcome Trust and Four Rivers Group that amount to $87 million in total (ResearchGate, 2017b).

The platform brings together individual user profiles, their publications, and room for discussion. After creating a profile, users share information about themselves (such as job position, education, projects and personal information), create a list of their publications and are invited to upload documents that are then accessible to others. Like Academia.edu, it is a peer-to-peer PDF sharing repository (Duffy & Pooley, 2017). Users can also follow other researchers and be followed by others. Finally, they receive personalised ads for jobs or products they might be interested in. It is reported that for some users ResearchGate created automatic profiles scraping publicly available information about users and publicly accessible documents in the PDF format (Van Noorden, 2014).

On the one hand, ResearchGate provides services to individuals, which are mainly targeted to their personal branding and promotion of their work. On the other hand, ResearchGate provides means for interaction among users and sharing their work. It motivates debate and encourages users to pose questions related to their work or to answer posted questions. Like Academia.edu it has most recently introduced a 'recommendation' function that makes it possible for other users to recommend someone else's work (ResearchGate, 2017a).

Academic social media platforms and big data

These three platforms have attracted millions of academic users by offering them use value. They collect large amounts of data on academics, their outputs, networks and institutions. While there are arguments that conceptualise social media and digital participation as something inherently new and different from social life beforehand, Marres (2017) convincingly argues that from a sociological perspective digital participation is not that different to older regimes

of media consumption; and that the uses and analyses of big data are merely an extension of social research traditions for knowing and intervening in society. Therefore, the question arises of how we can use these large data sets to continue with the study of globalisation, social relations and networks in higher education; and what are the methodological implications of doing so. In what follows, I will offer an overview of research on these matters across various social science disciplines.

Old wine in new bottles: methodological continuation in social enquiry

Academic social media platforms have attracted some attention among scholars, though not much within the field of education. The exception is a large body of work on technology in teaching and learning that I will leave out of the analysis as the focus of this paper is on academic social relations, governance and networks, rather than teaching and learning.

Academic social media have mostly been studied in the following disciplines: information science and technology, bibliometric and related studies, library and information science and media studies. In what follows, I present the findings of studies that specifically focus on one or more of the three case platforms in this chapter. Thus, I will be leaving out of the analysis studies on academics using other social media platforms, such as Twitter.

Increasing use of social media and the emergence of altmetrics

A number of studies have looked at academics' use of the Internet and how they have adapted to the ubiquity of digital technology (Gruzd et al., 2012; Mas-Bleda et al., 2014; Nicholas et al., 2015; Ortega, 2017; Procter, Williams, & Stewart, 2010; Rowlands et al., 2011; Van Noorden, 2014), and an overall scoping review has also been conducted (Kjellberg, Haider, & Sundin, 2016). There are few consistent trends found in these studies.

First, when academics consider where to publish their work, they still place highest value on the well-established channels of communication (i.e. academic journals). However, there is increasing use of the Internet and social media throughout the research cycle. Thus, academics see Web tools not as substitutes to traditional means of communication, but as having their own specific functions at particular stages of research. Moreover, researchers are broadly supportive towards using Web tools. Google Scholar seems to be most popular based on the findings in these studies, followed by ResearchGate and then Academia.edu.

Second, studies find that age is not an important factor in influencing the use of academic social media. Older academics are using social media equally frequently as their younger colleagues. However, academic status matters in that senior level academics tend to make less use of academic social media profiles than younger academics. Social media are primarily used for widening collaboration and there are signs that some researchers use social media to learn about research communities beyond their personal networks or to help them filter information. Some assume that this is because younger scholars have not yet developed strong personal networks and consequently find social media useful (e.g. Gruzd et al., 2012).

Finally, some academics reported that Web tools promote their work among people who might otherwise not hear of it. Many also report that Internet tools facilitate and promote international collaboration by helping with finding professional connections and maintaining existing ones. Procter et al (2010) found correlation between the use of Web tools and researchers' involvement in collaboration across institutions and countries. Some researchers also value the informality of communicating in these new ways.

This change in the format and distribution of academic publishing, as well as academic engagement in social media platforms, is potentially changing the way in which the quality of academic outputs can be evaluated. Traditional measures under the umbrella term of bibliometrics seem to be losing ground (Hoffmann, Lutz, & Meckel, 2016; Roemer & Borchardt, 2012; Yu et al., 2016), while we can notice the rise of so-called 'altmetrics' (alternative metrics) – Web based alternatives to measuring scientific impact. Altmetrics are

based on clearly defined social media platforms, which often provide free access to usage data through "Application Programming Interfaces" (APIs). This development has made it possible to analyse online usage of research outcomes independently of publishers and research impact can be analysed more broadly (Haustein et al., 2014). Thus, altmetrics allow for new maps of scholarly contribution, which are unprecedented in their subtlety, texture and detail (Priem, 2013).

Reputation and research evaluation

Although altmetrics builds on data from a variety of social media and other digital platforms, the focus here is on the three case platforms. A number of studies have looked at the reliability or effectiveness of these platforms and their constructed scores as measures of research quality (Martin-Martin et al., 2016; Mas-Bleda et al., 2014; Orduna-Malea, et al, 2017; Ortega, 2015c; Thelwall & Kousha, 2015, 2017a, 2017b; Yu et al., 2016). Most of these studies look at statistical correlation of the platforms' indicators with established measures of research quality at the individual, group, institutional or national levels.

Google Scholar Citation offers measures of research quality such as number of citations and h-index. ResearchGate created several analytical metrics, with the flagship performance indicator being its RG Score. It brings together bibliometrics and altmetrics, but is calculated by an undisclosed algorithm, which includes contributions to users' profiles, interactions with other members, and reputation among other members (Thelwall & Kousha, 2015). Academia.edu offers a variety of analytical metrics relating to user profiles, such as number of profile views and, as of 2017, AuthorRank and PaperRank.

Most studies have focused on the RG Score. Yu et al (2016) provide empirical evidence demonstrating that RG Score effectively measures institutional and individual researcher performance. Hoffman et al (2016) focused on the relational dynamics of networks on ResearchGate and found that the online publication resonance is highly correlated with the traditional offline metric (h-index) and also seniority network. Thelwall and Kousha (2015) find that ResearchGate statistics correlate moderately well with the five major

university rankings and suggest that ResearchGate broadly reflects the traditional distribution of academic capital. Ortega (2015) found that bibliometric indicators are stable across platforms and have high correlations, while social and usage indicators (such as number of followers) are platform-specific. Martin-Martin et al (2016) found that ResearchGate indicators present a high correlation to all the indicators from Google Scholar Citations. Although most scholars find that RG Score can be seen as a measure of research quality, it is also subject to critique. Kraker and Lex (2015), for example, claim that the score is not fully transparent and reproducible and changes in the score cannot be fully reconstructed.

Regarding Academia.edu metrics, Thelwall and Kousha (2014) found that traditional bibliometric measures did not correlate with any Academia.edu metrics on their sample. The authors believe this might be because of the range of informal scholarly activities that cannot be measured by bibliometric methods, or because more senior academics use the site less extensively.

Studying networks

Besides quantitatively measuring research performance and reputation, a final group of studies has analysed the data of academic social media to study forms of user behaviour (Hoffmann et al., 2016; Ortega, 2015a, 2015b, 2017, Thelwall & Kousha, 2014, 2015). They focus on various elements of population dynamics, such as how the networks grow on these platforms, differences in the use of platforms based on the users' discipline, gender, age, status, and so on.

Two studies found that user behaviour on ResearchGate and Academia.edu reaffirm established academic hierarchies, however, there is also a notable platform-specific dynamic that allows junior academics to build their academic networks (Hoffmann et al., 2016; Thelwall & Kousha, 2014). In studying the behaviour of academics from different countries and the possibility for the digital platforms to enable open access to research outputs, Thelwall and Kousha (2015) found that while academics from Brazil, India, and some other countries seem to be disproportionately taking advantage of ResearchGate to disseminate their work, academics in China, South

Korea, and Russia may be missing opportunities to use ResearchGate to maximize the academic impact of their publications.

Regarding disciplinary differences, it was found that Academia.edu is mostly populated by users from humanities and social sciences, while ResearchGate by users from biology and biomedicine (Ortega, 2015a). However, this trend is changing and users seem to be more and more representative of the 'real world' disciplinary distribution (Ortega, 2017). The growth in user numbers seems to be coming in waves emanating from specific institutions and countries (Ortega, 2015b). Users seem not to have several profiles across platforms and most only have one such profile (Ortega, 2015c).

There are no studies yet that use big data from these academic platforms to study academic networks and actors, social and/or power relations, and globalisation, and that go beyond bibliometric or library and information studies.

Theoretical and methodological aspects of studying academic social media

The studies using data from digital platforms' discussed so far applied a variety of approaches and methods, out of which most are quantitative. First, studies that analysed the increasing use of social media and the purposes for their use employed online questionnaires, interviews and focus groups (Gruzd et al., 2012; Procter et al., 2010; Rowlands et al., 2011). They targeted various samples and variables ranging from disciplines to institutions or countries.

A second group of studies analysed academic platforms for their potential use in altmetrics or for the potential of indicators measuring research quality (for example, Martin-Martin et al., 2016; Ortega, 2015c; Thelwall & Kousha, 2017b). These studies collected large data sets from the Internet, namely from the Academia.edu, Google Scholar Citations and ResearchGate platforms, using the API or SQL scripts, SocSciBot crawls of Web pages and other software programmes to crawl various Web pages and extract data. The data were then "cleaned," which means that the duplicates for user profiles were deleted, data about the same user profiles merged, and so on. After the "data cleaning," various statistical analyses were applied,

most often descriptive statistics, correlation analysis (the specific correlations that were applied depended on the sample and the research question), principal component analysis, various regression models, or also various models of counting. The most common type of analysis in this group is correlation between various ResearchGate or Academia.edu indicators with other more established indicators of research quality, such as citation index or h-index. These were extracted from either Google Scholar, Scopus or Web of Science. The idea behind these correlation analyses is to check the "reliability" of the indicators that the new platforms provide; like the RG Score, for example.

These studies contribute important knowledge about the new academic platforms and their indicators, but are generally uncritical towards the previously established metrics of research quality. While on the one hand some authors recognise the critique of measures like the citation index or the h-index (e.g. Martin-Martin et al., 2016), they do not tackle this critique, but take them as "the" measure of research quality against which the new indicators are tested. In this sense, they reproduce older measures of research quality as the 'gold standard'. Moreover, when analysing indicators at the level of universities or countries, correlation analyses are done between these indicators and university rankings, such as QS University Rankings, Times Higher Education World University Rankings, and Academic Ranking of World Universities (ARWU). Again, this means taking the rankings as reliable and true measures of the research quality of various universities or countries, while we know that there are many problems with university rankings (Hazelkorn, 2009).

A third group of studies focused on behaviour of user networks and applied similar methods of data collection and analyses to the second group, but the research questions were different (for example, Ortega, 2015a; Thelwall & Kousha, 2017a). In this group of studies, networks of disciplines or countries were tested and correlation analyses between these different groups were conducted. Therefore, the types of statistical analyses are similar to the previous group of studies, but the variables for correlation analyses are different. Besides correlation analyses, descriptive statistics of various models of counting are performed, such as calculating compound annual growth rate. In this group of studies, we most often learn about differences in user

behaviour based on discipline, gender, academic rank or age for various samples as discussed before.

While studies in all of the three groups provide important empirical data and analyses, most of them are not explicit in their theoretical framing and normally do not go much further than providing data. A common thread in all of these studies is that there is not much use of theory to conceptualise the study or to interpret the analysis and findings (Kjellberg, Haider, & Sundin, 2016).

Therefore, what is missing are methodologies that bring together the quantitative analyses that are possible with large data sets extracted from the academic platforms and the theoretical interpretive frameworks of education studies and related disciplines like sociology of education or political economy of education. In other words, there is room for empirical studies that make use of large digital data sets to theorise the insights with specific interpretive frameworks. This would make it possible to interpret academic platform use and user behaviour in the light of the knowledge economy and global power relations, and allow the possibility to study academic platforms' power to construct new and different relations between academics, universities and countries.

New wine in new bottles: Social media platforms constituting new social relations

What I have showed so far is the uptake of the whole of the research cycle and the production of knowledge and academic life/communication onto new digital platforms (such as: discussing research ideas, finding research partners, presenting and sharing research data and its analysis, drafting academic papers, collecting comments on papers, publishing final papers, evaluating papers). The question that now follows is whether these platforms structure or re-structure particular social relations. In other words, what particular forms of social relations does the infrastructural design of these platforms enable and what forms of social relations does it disable. I will tackle this question in three stages with which I aim to address the methodological shortcomings identified in the previous section by

extending the analysis of the academic platforms into more critical territories.

Processes of quantification and individualisation

The analysed platforms attempt to reproduce the production of knowledge on the Web, or move it to the Web. Researchers and institutions are identified, who are then made users of platforms and who populate their profiles with data. They are encouraged to share their work products, to read the work of others, comment and make recommendations. Moreover, new features such as "recommendations" in Academia.edu and its "Editors program" attempt to disrupt the established peer review system. Users are motivated to communicate with each other and engage in scholarly debate. The platforms' core aim is to lubricate connections and networks. Finally, the platforms provide numerical feedback in the form of analytics about the user profiles (see Table 1).

So far so good. It seems as if the research process is intact and that academics reap benefits by engaging with these platforms. However, the process of the academic knowledge production is now quantified and every user activity gets allocated a number. Second, the process becomes "platform ready" (Helmond, 2015). This means that other webpages on the Internet structure their data so that they can be found and used by these three platforms, or that these three platforms are integrated into other webpages so that users can share the information across platforms and websites with the click of a button. For example, if authors or institutions want particular publications and documents to be found by Google Scholar, the files need to be appropriately assigned metadata.

The platforms not only sort the gathered data, but also use it in their algorithms. In the three case platforms, one of the important uses of data is for the production of impressively detailed analytics (see Table 1).

Table 1. Overview of analytics offered by the three digital platforms.

Academia.edu
Profile views: the number of times the profile has been viewed. **Paper views:** the amount of times all of user's papers have been viewed. **Unique visitors:** how many individual people have visited user's papers or profile. **Profile visitors:** offers information about which part of Academia.edu profile the visitor went to, last page the visitor was at before coming to the Academia.edu profile, the city and country in which the visitor is located, and Premium subscribers also receive information about which university visitors come from and what is their role. **Keywords:** are the search queries that people have entered on Google and other search engines to find user's work on Academia.edu. **AuthorRank and PaperRank:** this is a new feature (as of 2017). The PaperRank of a paper is a function of the number of recommendations the paper has received, weighted by the AuthorRanks of the recommenders. The AuthorRank is a function of the PaperRanks of the papers on the user's profile. It is impossible to have AuthorRank without PaperRank, and it's impossible to have PaperRank without being Recommended. **Recommendations:** (this is a new feature as of 2017) Papers can be recommended as worthwhile contributions to the body of literature in their field. Access to this feature is available to senior academics through the Editor Program. **Analytics open to Premium subscription**: mentions, readers, enhanced analytics (and advanced search of papers).
Google Scholar Citations
Citations: Number of citations of all publications. Computed for citations from all years, and citations since 2012. **h-index:** The largest number h such that h publications have at least h citations. Computed for citations from all years, and citations since 2012. **i10 index:** Number of publications with at least 10 citations. Computed for citations from all years, and citations since 2012. **Cited by:** The total number of publications in which the particular paper was cited.
ResearchGate
Total publications: The number of publications listed by an academic in their profile. **Total impact points:** The cumulative journal impact factors of the publications of an academic. **Reads:** Total number of times an author's contributions to ResearchGate have been visualized, recently combined with the total number of downloads recorded by ResearchGate for the full-text articles uploaded to author profiles. **RG Score:** This is a number for each academic and institution calculated by ResearchGate using an algorithm that is not fully disclosed but which is based on contributions to members' ResearchGate profiles, interactions with other members, and reputation among other members. **Citations:** Total number of citations to the documents uploaded to the profile. ResearchGate generates its own citation database. **Impact points:** Sum of the JCR impact factors of the journals where the author has published articles. **Profile views:** Number of times the author's profile has been visited. **Following:** Number of ResearchGate users the author follows. **Followers:** Number of ResearchGate users who follow the author.

Sources: On Google Scholar Citations: (Martin-Martin et al., 2016). On ResearchGate: (Martin-Martin et al., 2016; Thelwall & Kousha, 2015). On Academia.edu: (Academia.edu, 2017c, 2017d, 2017e, 2017g).

As the research process and scholarly communication are made 'platform-ready' and assigned a number, this has effects on individuals. The platforms provide detailed data about a user's profile

that become increasingly refined and even more detailed over time. Structurally, this creates constant self-monitoring and self-surveillance. By analysing Academia.edu, Duffy and Pooley (2017) argue that the platform becomes fixated on analytics and, as a result, reinforces a culture of incessant self-monitoring. It is safe to extrapolate this to the other two platforms under consideration here. These new forms and ways of measuring and evaluating oneself lead to new forms and ways of comparison and, consequently, competition. The kind of competition and restructuring of social relations that I turn to next, are particularly led by network effects enabled by digital platforms. Thus the following processes refer more to Academia.edu and ResearchGate than to Google Scholar Citations.

Processes of enhancing competition

As already mentioned, the analytics that users see about other users are more limited than about oneself, i.e. for other users they can see h-index, number of citations, RG score and AuthorRank. With the exception of Google Scholar Citations, the analytics of the other two platforms take social media metrics and marry them to academic measures of quality. Consequently, academics are structurally expected to compete not only with academic outputs and their scores as traditionally measured with citation indexes, but also with social media activity, because this will improve their scores on those platforms. Thelwall and Kousha find this in their study of ResearchGate: "[RG Score…] is based on contributions to members' ResearchGate profiles, interactions with other members, and reputation among other members. ... In addition, its activity component gives a large bias towards academics and institutions that employ ResearchGate the most, making it a hybrid scholarly achievements and site use indicator" (Thelwall & Kousha, 2015, p.880). Moreover, Orduna-Malea et al (2017) found that authors achieve high RG Scores by social and collaborative activity on the platform, particularly by answering questions, which is one of the services on ResearchGate. They also found that it is difficult for authors to reach high scores from publications alone.

Some of these platforms suggest to users how to improve their statistics, which is normally by contributing even more content and engaging with more people and producing even more data. At the same time, the production of data and academics' responses to platforms' suggestions give the companies more opportunities to refine their algorithms for analytics purposes. Thus, for companies it is not only about big data, but the use of big data and the creation of algorithms to intervene in the social world and to structure social relations (Fourcade & Healy, 2017).

Social media turn users into prosumers (Ritzer, Dean, & Jurgenson, 2012) in that while they are consuming the service offered by the social media, they are at the same time producing it. Academic social media, therefore, include the invisible academic labour of traditional academic communication, including peer review, together with invisible academic labour producing the social media content. Ironically, if academics engage with social media and take their analytics seriously, then they start competing not only in research quality (however it is measured), but in "presumption" of the social media platforms. These scores are promoted as measures of research quality, but they in fact include social media activity. It seems that academic competition in research quality is being reframed into academic competition in skilful use of academic social media.

Authors recognise that digitalisation and platformization often goes hand in hand with privatization and monetization, although they are not the same processes (Marres, 2017). While the social media companies legitimize their activities according to the discourses of "open science" and democratization of knowledge production and dissemination, at the same time there are instances of monetizing the data and the hidden un-paid labour behind the platforms. Such is the case of Academia.edu and its attempts to monetise academics' appetites for competition and standing in their fields, which ended in academic uproar. Bond (2017) made a public call to academics to delete their Academia.edu accounts for several reasons. One of them was the company's actions when they sent emails to scholars asking them if they would pay a "small fee" for getting their papers recommended on the platform. As this was not well received, Academia.edu moved to charging subscription to its "premium feature".

Regardless of whether social media are free to use or not, the statistics that they produce represent new forms of academic competition in size, scope, reach and temporality. Most importantly, they are enforced by the platforms' infrastructure, which has effects on wider social relations among scholars.

Processes of restructuring social relations

The design of academic platforms structure social relations in many ways. First, by classifying who gets to do what, as in the case of Academia.edu and its relatively new Editor Program, which replicates the traditional peer review process. Users are asked to make an application to get rights to recommend papers. The criteria seem scholarly, namely to "have a doctorate or professorship in the field; have some experience publishing papers; have some experience peer reviewing papers." (Academia.edu, 2017f) However, the scores of papers are based on who recommends them and their personal scores. These personal scores of the users who make recommendations are impacted again by that person's contribution to the platform in combination with standing in the field. Thus there is a never-ending snowball of metrics that assign worth to users and determine how are they related to whom. Academia.edu predicts that it will incorporate an increasing number of academics into the Editor Program and that recommending papers on the platform will become their standard practice (Academia.edu, 2017f).

Platforms can also structure new forms of academic hierarchy. As Bond (2017) comments, Academia.edu's Premium feature "allows users to get special data analytics about who is reading their papers, including the "role" (i.e. the rank) of the person looking at their work. Emails even go out to users letting them know the percentile (a top 4% scholar!) of the person downloading their work. Are we supposed to somehow value that a full professor looked at our work over, say, an adjunct? The new feature is academic class politics to a new level and it only promotes the further stratification of the academy" (Bond, 2017).

Duffy and Pooley (2017) examine Academia.edu and analyse the deeper forms of academic hierarchy that the platforms' algorithms are

now promoting. They find that the platform encourages the specific kind of subjectivity and logic of self-branding and carefully curated self-promotion. This potentially brings new elements into academic social relations. The academic world was always competitive, however, the detailed analytics of self-monitoring brings about new dimensions of individuality.

Another important structuring of social relations is that the platforms focus on individualisation in providing suggestions on whom to read, whom to connect with, whom to follow, and also offering particular advertisements for jobs or services. Fourcade and Healy (2017) argue that this is the new dimension of social stratification. The platforms and their algorithms assign the individual a certain role, position them in a certain place in a certain network and, based on that, they get particular opportunities and offers, while others do not. In this sense, the social relations are neither spontaneous nor orchestrated through power relations between actors like before. Rather, this is a new dimension of data and algorithms.

All in all, these academic digital platforms make possible new forms of disseminating knowledge by enabling comments and debate as academic papers are developed. They provide new modes of evaluating research with altmetrics that include Web page views, reader votes, comments and the like. Moreover, new forms of filtering mean personalisation and custom pre-filtering of relevant research output. Finally, the hiring and reward system of scholars are expected to change so that scholars can be hired and paid based on the impact of their diverse products, which is measured with Web metrics (Priem, 2013). Academic platforms have the potential to become classifiers of academics, their outputs and their relations as well as of university departments, universities or countries.

Conclusion

This chapter has explored the complex reality of the digitalization and platformization of higher education by focusing on academics and the research processes they perform. My entry point was three digital platforms targeting academics that collect, analyse and use large amounts of data, contributing to the "big data" phenomenon.

My first aim was to discuss what we have learned about higher education, academics, their relations and networks by analysing the big data produced by these platforms. I found that while education scholars have not engaged with analysis of big data from the three particular digital platforms, scholars of other disciplines have, particularly in bibliometrics and library and information studies. Most work has analysed altmetrics and how the platforms' scores correlate with established measures of quality, such as citation indicator or h-index. What seems to be missing is marrying the empirical data analysis with theoretical frameworks of education studies or related disciplines. There is significant potential for exploring web-based academic actor-networks and for studying international power relations in the knowledge economy at individual, institutional or state levels using big data generated by these platforms.

My second aim was to scrutinise how the three academic platforms are structuring academic social relations. I found important effects of academic digital platforms as they provide use value to academics. I identified and described three processes: quantification and individualization; enhancing competition; and re-structuring academic social relations. As academics engage with these digital platforms, they create data, which are instantly processed with algorithms that provide immediate individual or aggregate analytics. This process quantifies the academic research process in possibly new ways and structures subjectivity and individuality differently. It enhances different types of academic valuation in which competition in research quality now includes social media activity. Algorithms of social media platforms also become structuring actors of academic relations as they promote particular people, papers and jobs, and not others.

It remains to be seen what the future effects of academic social media and digital platforms will be. This chapter has contributed to the debate by illuminating the need for methodological innovation to study these important and unfolding dynamics.

References

Academia.edu. (2017a). About. Retrieved August 24, 2017, from http://www.academia.edu/about

Academia.edu. (2017b). Academia.edu. Retrieved September 1, 2017, from https://www.academia.edu/

Academia.edu. (2017c). Academia Premium Features. Retrieved September 15, 2017, from http://support.academia.edu/customer/en/portal/topics/896070-academia-premium-features/articles

Academia.edu. (2017d). Analytics Overview. Retrieved September 15, 2017, from http://support.academia.edu/customer/en/portal/articles/2313235-analytics-overview

Academia.edu. (2017e). Recommendations. Retrieved September 15, 2017, from http://support.academia.edu/customer/en/portal/articles/2130222-recommendations

Academia.edu. (2017f). The Academia.edu Editor Program. Retrieved September 15, 2017, from https://www.academia.edu/editor-program

Academia.edu. (2017g). What are AuthorRank and PaperRank? Retrieved September 15, 2017, from http://support.academia.edu/customer/en/portal/articles/2201342-what-are-authorrank-and-paperrank-

Bond, S. (2017, January 23). Dear Scholars, Delete Your Account At Academia. *Forbes*. Retrieved from https://www.forbes.com/sites/drsarahbond/2017/01/23/dear-scholars-delete-your-account-at-academia-edu/#738fbd4c2d62

Duffy, B. E., & Pooley, J. D. (2017). “Facebook for Academics”: The Convergence of Self-Branding and Social Media Logic on Academia.edu. *Social Media and Society*, 3(1), 1–11.

Fourcade, M., & Healy, K. (2017). Seeing like a market. *Socio-Economic Review-Economic Review*, 15(1), 9–29.

Google Scholar. (2017a). Google Scholar. Retrieved September 1, 2017, from https://scholar.google.co.uk/intl/en/scholar/about.html

Google Scholar. (2017b). Google Scholar Profile. Retrieved September 1, 2017, from https://scholar.google.co.uk/

Gruzd, A., Staves, K., & Wilk, A. (2012). Connected scholars: Examining the role of social media in research practices of faculty using the UTAUT model. *Computers in Human Behavior,* 28(6), 2340–2350.

Haustein, S., Peters, I., Bar-Ilan, J., Priem, J., Shema, H., & Terliesner, J. (2014). Coverage and adoption of altmetrics sources in the bibliometric community. Scientometrics, 101(2), 1145–1163.

Hazelkorn, E. (2009). Rankings and the battle for world-class excellence: Institutional Strategies and Policy Choices. *Higher Education Management and Policy*, 21(1),

Helmond, A. (2015). The Platformization of the Web: Making Web Data Platform Ready. *Social Media and Society*, 1(2), 1–11.

Hoffmann, C. P., Lutz, C., & Meckel, M. (2016). A Relational Altmetric? Network Centrality on ResearchGate as an Indicator of Scientific Impact. J*ournal of the Association for Information Science and Technology,* 67(4), 765–775.

Kjellberg, S., Haider, J., & Sundin, O. (2016). Researchers’ use of social network sites: A scoping review. *Library and Information Science Research*, 38(3), 224–234.

Kraker, P., & Lex, E. (2015). *A Critical Look at the ResearchGate Score as a Measure of Scientific Reputation*. In Proceedings of the Quantifying and Analysing Scholarly Communication on the Web workshop (ASCW'15), Web Science Conference.

Liu, Z. (2003). Trends in transforming scholarly communication and their implications. Information Processing and Management, 39(6), 889–898.

Marres, N. (2017). *Digital sociology: the reinvention of social research*. Cambridge: Polity Press.

Martin-Martin, A., Orduna-Malea, E., Ayllon, J. M., & Lopez-Cozar, E. D. (2016). The counting house, measuring those who count: Presence of Bibliometrics, Scientometrics, Informetrics, Webometrics and Altmetrics in the Google Scholar Citations, ResearcherID, ResearchGate, Mendeley & Twitter (No. 21). https://doi.org/http://doi.org/10.13140/RG.2.1.4814.4402

Mas-Bleda, A., Thelwall, M., Kousha, K., & Aguillo, I. F. (2014). Do highly cited researchers successfully use the social web? *Scientometrics*, 101(1), 337–356.

Nicholas, D., Herman, E., Jamali, H., Rodriguez-Bravo, B., Boukacem-Zeghmouri, C., Dobrowolski, T., & Pouchot, S. (2015). New ways of building, showcasing, and measuring scholarly reputation. *Learned Publishing*, 28(3), 169–183.

Orduna-Malea, E., Martín-Martín, A., Thelwall, M., & Delgado López-Cózar, E. (2017). Do ResearchGate Scores create ghost academic reputations? *Scientometrics*, 112(1), 443–460.
Ortega, J. L. (2015a). Disciplinary differences in the use of academic social networking sites. *Online Information Review,* 39(4), 520–536.

Ortega, J. L. (2015b). How is an academic social site populated? A demographic study of Google Scholar Citations population. *Scientometrics*, 104(1), 1–18.

Ortega, J. L. (2015c). Relationship between altmetric and bibliometric indicators across academic social sites: The case of CSIC's members. Journal of *Informetrics*, 9(1), 39–49.

Ortega, J. L. (2017). Toward a homogenization of academic social sites: a longitudinal study of profiles in Academia.edu, Google Scholar Citations and ResearchGate. *Online Information Review*, 41(6), 812–825.

Oxford Dictionary. (2016). Definition of social media in English. Retrieved January 20, 2016, from http://www.oxforddictionaries.com/definition/english/social-media?q=social+media

Price, R. (2017). About Richard. Retrieved September 1, 2017, from http://www.richardprice.io/about

Priem, J. (2013). Scholarship: Beyond the paper. *Nature*, 495(7442), 437–440.

Procter, R., Williams, R., & Stewart, J. (2010). If You Build It, Will They Come? How Researchers Perceive and Use Web 2.0. London: Research Information Network. Retrieved from http://www.rin.ac.uk/system/files/attachments/web_2.0_screen.pdf

ResearchGate. (2017a). About Us. Retrieved August 24, 2017, from https://www.researchgate.net/about

ResearchGate. (2017b). About Us: ResearchGate Factsheet. Retrieved September 1, 2017, from https://www.researchgate.net/aboutus.AboutUsPress.downloadFile.html?name=rg_fact_sheet.pdf

Ritzer, G., Dean, P., & Jurgenson, N. (2012). The Coming of Age of the Prosumer. *American Behavioral Scientist*, 56(4), 379–398.

Roemer, R. C., & Borchardt, R. (2012). From bibliometrics to altmetrics: A changing scholarly landscape. *College & Research Libraries News,* 73(10), 596–600.

Rowlands, I., Nicholas, D., Russell, B., Canty, N., & Watkinson, A. (2011). Social media use in the research workflow. Learned Publishing, 24(3), 183–195.

Scott, M. (2017). A Facebook-Style Shift in How Science Is Shared. The New York Times. Retrieved from https://www.nytimes.com/2017/02/28/technology/science-research-researchgate-gates-goldman.html

Shead, S. (2017). Goldman Sachs and Bill Gates quietly invested $52 million in a social network for scientists. *Business Insider Deutschland*. Retrieved from http://www.businessinsider.de/researchgate-raised-52-million-goldman-bill-gates-2017-2?r=UK&IR=T

Thelwall, M., & Kousha, K. (2014). Academia.edu: Social Network or Academic Network? *Journal of the Association for Information Science and Technology,* 65(4), 721–731.

Thelwall, M., & Kousha, K. (2015). ResearchGate: Disseminating, Communicating, and Measuring Scholarship? *Journal of the Association for Information Science and Technology*, 66(5), 876–889.

Thelwall, M., & Kousha, K. (2017a). ResearchGate Articles: Age, Discipline, Audience Size and Impact. *Journal of the Association for Information Science and Technology,* 68(2), 468–479.

Thelwall, M., & Kousha, K. (2017b). ResearchGate versus Google Scholar: Which finds more early citations? *Scientometrics*, 112(2), 1125–1131.

UNESCO. (2015). *UNESCO Science Report: Towards 2030 Executive Summary*. Paris: United Nations Educational, Scientific and Cultural Organization.

Van Noorden, R. (2014). Scientists and the Social Network. *Nature*, 512 (12 August), 126–129.

Yu, M. C., Wu, Y. C. J., Alhalabi, W., Kao, H. Y., & Wu, W. H. (2016). ResearchGate: An effective altmetric indicator for active researchers? Computers in Human Behavior, 55, 1001–1006.