North Perspectives for a Better South?
Big Data and the Global South in Big Data & Society

Guilherme Cavalcante Silva

Universidade Estadual de Campinas - guilhermecavalcantesilva@outlook.com

Abstract

Over the last few years, data studies within Social Sciences watched a growth in the number of researches highlighting the need for more proficuous participation from the Global South in the debates of the field. The lack of Southern voices in the academic scholarship on the one hand, and of recognition of the importance and autonomy of its local data practices, such as those from indigenous data movements, on the other, had been decisive in establishing a Big Data in the South agenda. This paper displays an analytical mapping of 131 articles published from 2014-2016 in Big Data & Society (BD&S), a leading journal acknowledged for its pioneering promotion of Big Data research among social scientists. Its goal is to provide an overview of the way data practices are approached in BD&S papers concerning its geopolitical instance. It argues that there is a tendency to generalise data practices overlooking the specific consequences of Big Data in Southern contexts because of an almost exclusive presence of Euroamerican perspectives in the journal. This paper argues that this happens as a result of an epistemological asymmetry that pervades Social Sciences.

Keywords: Global South, Big Data, Big Data & Society, Social Sciences.

Perspetivas do Norte para um Sul melhor?
Big Data e o Sul Global na Big Data e Sociedade

Sumário

Nos últimos anos, os estudos de dados nas Ciências Sociais observaram um crescimento no número de pesquisas destacando a necessidade de uma participação mais profícua nos debates da área por parte do Sul Global. A falta de vozes do Sul na produção de conhecimento académica por um lado, e de reconhecimento da importância e autonomia das suas práticas de dados locais como os movimentos de dados indígenas, por outro, foram decisivos no estabelecimento de um Big Data na agenda do Sul. Este artigo apresenta um mapeamento analítico de 131 artigos publicados de 2014 a 2016 na Big Data & Society (BD&S), uma revista líder reconhecida pela promoção pioneira da pesquisa em Big Data entre cientistas sociais. O seu objetivo é fornecer uma visão geral da maneira como as práticas de dados são abor-

dadas nos artigos da BD&S em relação à sua instância geopolítica. Argumenta que há uma tendência para generalizar práticas de dados negligenciando as consequências específicas do Big Data em contextos do Sul devido a uma presença quase exclusiva das perspetivas euro-americanas na revista. Este artigo defende que isto acontece como resultado de uma assimetria epistemológica que permeia as Ciências Sociais.

Palavras-chave: Sul Global, Big Data, Big Data e a Sociedade, Ciências Sociais.

INTRODUCTION

The term Big Data has become increasingly popular in academic researches in multiple fields, as well as in news reports and business reports. Although it has been circulating in informational environments since the end of the 20th century, it is only in the decade of 2010 that Big Data gains popularity in scientific productions[1] and becomes a catchphrase (Tomaz & Silva, 2018). Within Social Sciences, the same phenomenon happens. An example of that is the creation of the journal Big Data & Society (BD&S) in 2014, explicitly focused on analysing "Big Data practices [...] while also reflecting on the consequences for how societies are represented (epistemologies), realised (ontologies) and governed (politics)"[2] in dialogue with concerns of the Social Sciences.

Considering the infancy of Big Data research, this work carried out an analytical mapping on Big Data research, especially on those perspectives of researchers from the many different Social Sciences. This study takes as its study object the articles published in BD&S between 2014 and 2016, totalling 131 articles. The goal was to understand the formation of the trends, venues, and canons that ground BD&S approaches, both through a quantitative analytical mapping and through ethnographic accounts of participation in events promoted by the BD&S community. Thus, the idea was to understand how the venues and scientific communities leading BD&S modulate the viewpoints that appear there.

---

1 Between 2012 and 2017, for example, the publication rate for "Big Data" was quintupled, according to data from the Web of Science. Available at https://wcs.webofknowledge.com/RA/analyze.do?product=WOS&SID=8E7jMfgE17MhNfzZc9B&field=PY_PublicationYear_PublicationYear_en&yearSort=true

2 About the Journal, Big Data & Society. Available at http://bigdatasoc.blogspot.com.br/p/big-data-and-society.html

This paper presents a partial view of such study, especially its quantitative aspect, paying particular attention to the question of the geographical distribution of voices that circulate in the discussion on BD&S. This concern arises in a context of growing interest over the participation of Global South actors in formulating the directions of areas related to the Social Sciences. Many see a neglect of such a question as resulting in a practice of generalising about contexts of such regions "rather than including voices from within the continent[s]" (Ganter & Ortega, p. 68). In the context of Big Data studies, this usually leads to "hyperbolic narratives of the 'big data revolution'" (Milan & Treré, 2019, p. 320) which, however, does not present itself as a revolutionary and 'effective' process in the experience of marginalised groups from the Global South[3], as attested for example in the state surveillance practices against indigenous peoples around the Globe (Kukutai & Taylor, 2016, Mann & Daly, 2018).

The purpose of this work is to identify from the analysis of the articles of BD&S, the asymmetric distribution of voices within BD&S' scientific communities and evaluate the ways in that such distribution affects the scenario of this research field. It also seeks to understand how that scenario informs the impact of Big Data in the Global South. The argument here is that such hegemony of Northern perspectives in the papers published on BD&S can result in a constant generalisation of the implications of Big Data, especially when it comes to Big Data developments in Southern contexts.

That is true since "the majority of the world's population today resides outside the West", even if the debate is still framed "'by means of Western' concerns, contexts, user behaviour patterns, and conceptual frameworks" (Milan & Treré, 2019, p. 320). We conclude that there is an urgent need to overcome the historical asymmetry that is actualising itself in Big Data. This work argues for a flourishing of critical perspectives concerning a "universal Big Data" view typical of liberal realities. To this end, we must recognise that the resonation of BD&S discourses in the communities in which it resonates plays a significant role in the absence of awareness of this asymmetry. That involves the language in which the papers are written and the almost exclusively Euroamerican theoretical canon that finds echo there.

The remainder of the paper is structured as follows. Firstly, the paper introduces the researches on Big Data, especially those from the Social Sciences to situate the most influential views on the subject currently circulating. Secondly, the article then

---

3 This paper follows the definition gave by Stefania Milan and Emiliano Treré (2019, p. 321) regarding what is the Global South. Beyond the obvious geographical connotation, naming where the marginalized are to be generally found, such a concept is used to identify "a plural entity subsuming also the different, the underprivileged, the alternative, the resistant, the invisible, and the subversive".

presents the issues of the Global North-South relationship in Big Data studies and makes its case for a de-westernization of Social Sciences researches in general – and in Big Data in particular. Thirdly, it will present a summary of the mapping on BD&S with a focus on the geopolitical issues surrounding the research production at the Big Data-Social Sciences interface. Finally, the paper briefly discusses the results of the analytical mapping by placing them in the context of BD&S circulation venues. That way, it seeks to understand how the operation of BD&S ends up 'shutting the door' for the recognition of North-South asymmetry.

This effort intends to contribute to the newly formed "Big Data from the South" agenda, led by theoreticians like Stefania Milan, Emiliano Treré, Payal Arora, among others. It is important to note that the focus of this paper is on BD&S and therefore covers a limited scope for analysing the relationship between Big Data and Social Sciences. Further research could contemplate other publication and bring more data on the subject.

BIG DATA RESEARCH: AN OVERVIEW FROM SOCIAL SCIENCES

Big Data research has witnessed an increasing rate in the number of papers around Big Data encompassing areas as diverse as Computing, Economics, Public Health, Education and Communication. The increase in the number of academic events[4], research groups, and formation programs[5] around the topic reveals the same reality.

Despite the great commotion around Big Data, there are few initiatives to this day devoted to mapping the main approaches in these researches. Some of these surveys encompass questions such as the state of the empirical research on Big Data (Wienhofen, Roman, & Mathisen, 2015), Big Data's definitions within Management (Ylijoki & Porras, 2016) and Digital Humanities (Kaplan, 2015). As far as Big Data studies in the Social Sciences are concerned, two papers are especially useful for understanding the paths walked so far. The first is an analytical mapping implemented by Jan Youtie, Alan L. Porter, & Ying Huang (2016) to evaluate the distribution of research interests within Big Data research in Social Sciences. The mapping comprised a total of 488 articles retrieved from the Web of Science (WoS) database. The other is also a

---

4 At websites dedicated to the report of CfP for scientific events, Big Data is among the most common conference themes. See for example https://www.papercrowd.com/conferences/search

5 Examples such as the one from *Data Diplomacy*, that involves academic institutions from New Zealand, Australia and England, investigating the role of democracy in what concerns data sharing; and the Data Institute, based in the University of San Francisco (US) and devoted to the formation of data scientists from all over the world and from multiple disciplines, demonstrate the transdisciplinary character of these efforts.

bibliometric endeavour produced by Jacky Akoka, Isabelle Comyn-Wattiaz, & Nabil Laoufi (2017) using a sample of 1843 articles hosted in the ScienceDirect database between 2013 and 2014. Their goal was to outline Big Data research in Social Sciences in comparison to other fields.

One of Youtie et al. (2016) first findings is that Big Data research in the sciences had a boom from the beginning of the 2010s. The year 2012, more specifically, marks the beginning of an exponential growth of interest in Big Data among social scientists, which has continued to grow to this day (Youtie et al., 2016). That was the year of the publication of the most cited work on Big Data in the field of Social Sciences: the article "Critical Questions for Big Data", authored by danah boyd[6] and Kate Crawford (2012) and published in Information, Communication & Society. From the year 2012 to 2017, the rate of publication on "Big Data" multiplied five times.
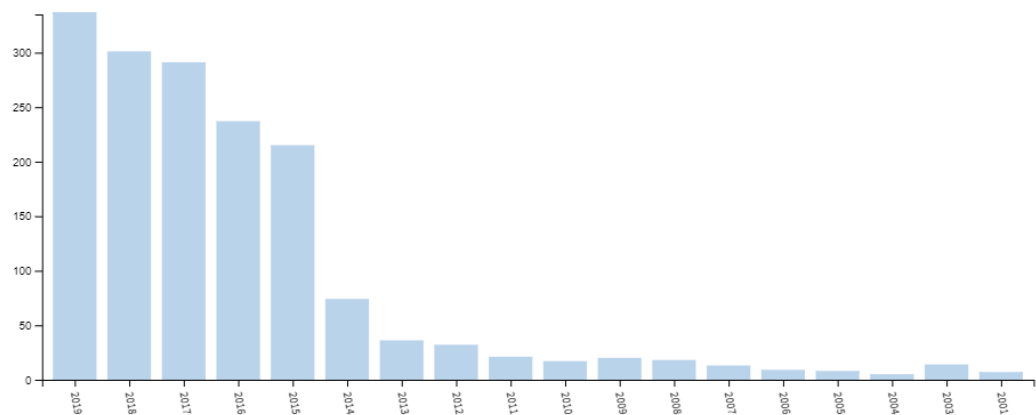

Figure 1. Number of published papers about "Big Data" in the fields of "Communication", "Arts", and "Humanities and Social Sciences" included in Web of Science's database between 2000 and 2019. Data updated until November 2019. Source: Web of Science Result Analysis. Available at https://wcs.webofknowledge.com/RA/analyze.do?product=WOS&SID=7FtwzFFyCvo7kGbTf1H&field=TASCA_JCRCategories_JCRCategories_en&yearSort=false

Youtie et al. (2016) also identified as central articulators of research interests in the Big Data-Social Sciences interface topics such as Internet and Society, Privacy Studies, Sociology of Science, Big Data and Medicine, Geolocalization, Decision Making,

6 boyd went to US court in 2011 and earned the right to be able to spell her first and last name in lower case. More details on the reasons for the change are available at http://www.danah.org/name.html

Business Impacts, and Analytics/Software. However, recalling the scope of this paper, it calls attention to the fact that researchers affiliated to US or British educational institutions authored 81% of the 488 articles.

In their mapping showing a broad picture of Big Data research in over 24 different disciplines, Akoka et al. (2017) provide an overview of Big Data research in Social Sciences as well. Big Data research in Social Sciences still falls short of other fields such as Computer Science and Engineering, both representing more than half of the total of articles on Big Data from 2000 to 2016. Despite that, Social Sciences stay as the fourth discipline with most publications on the list (Akoka et al., 2017).

Like Youtie et al. (2016), Akoka et al. (2017) point to a hegemonic presence of North American and European research institutions in Big Data research[7]. Figure 2 below illustrates such a discrepancy between Euroamerican participation and that of institutions, authors and movements of the Global South. It is an attempt to express the balance of power in the geopolitics of scientific, technological and symbolic production on Big Data.
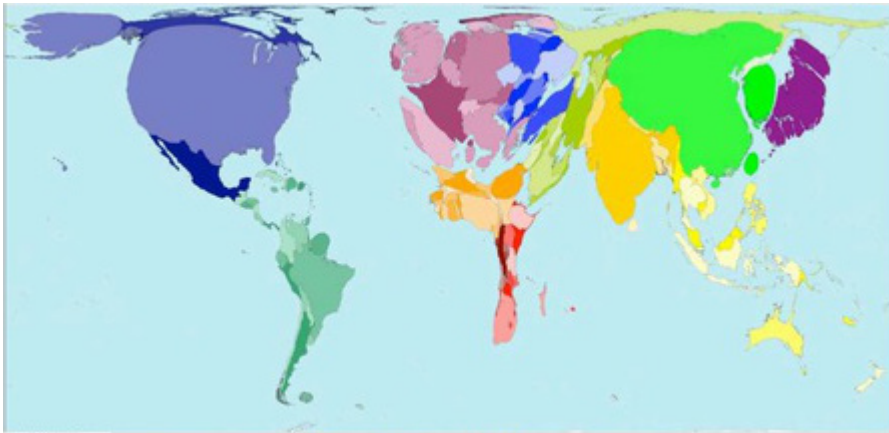

Figure 2. Geopolitical map of Big Data's global balance. Produced by computational scientist Stéphane Grumbach, from the Institut National de Recherche en Informatique et en Automatique (Inria / France). Source: "Big Data? The Global Imbalance". Available at http://www.in2p3.fr/actions/formation/Info13/lift-27sept-121001065613-phpapp01.pdf

---

7 Exception must be made here to the important participation of Chinese universities in the promotion of discussions on Big Data at a global level. This is noticeable, for example, in the number of international transdisciplinary journals under the leadership of Chinese researchers and institutions, such as the *International Journal of Big Data Intelligence* (National Chung Cheng University), *Big Data and Cognitive Computing* (Huazhong University of Science and Technology) , *Big Data Research* (Zhejiang University) and the *Open Journal of Big Data* (Xi'an Jiaotong Liverpool University).

In short, although recent Big Data studies within Social Sciences are on a steady rise, to the point where some works may already be classified as foundational (Iliadis & Russo, 2016). The consolidation of several lines of research, methodologies and different study objects also points to a growth in Big Data research in Social Sciences. However, this increasing volume of publications happens at the expense of greater participation of voices and ideas from the Global South, exposing the way Big Data affect the margins. Most importantly, it occurs at the cost of the recognition of the reality stated in Figure 2.

Despite providing an outlook on the Big Data research in Social Sciences and identifying this Global North/South divide in their surveys, Akoka et al. (2017) and Youtie et al. (2016) perform only a descriptive task in their papers. Their paper, for example, neglect showing how the papers analysed use their references or the reasons and consequences of the Euroamerican hegemony in Big Data research in Social Sciences. This question, however, should not be treated as being irrelevant, since it impacts the formulation of concepts, theories, and methods for studying Big Data in contemporaneity (Milan & Treré, 2019).

BIG DATA AND THE GLOBAL NORTH/SOUTH DIVIDE

One of the many examples of the importance of recognising the Global North/South divide is given by the Brazilian sociologist Francisco de Oliveira. In his classic "Critique to a Dualistic Reason", Oliveira (2003) pinpointed what he saw as a significant flaw in Latin American socioeconomic thought when discussing the specificity of the capitalist development in Brazil. On the one hand, the conventional economic analyses evaluated the Brazilian reality in terms of a deep-rooted inequality, leaving beyond question, on the other hand, the structure of a traditional western capitalist thought that appealed to dualities such as 'progress' and 'underdevelopment', 'modernity' and 'traditionalism'.

> While they denounced the miserable living conditions inflicted upon a large portion of the Latin American population, their theoretical and analytical schemes tied themselves to discussions around the relation between product-capital, propensity to save or invest, the marginal efficiency of capital, economies of scale, and size of the market. That led them, unwittingly, to construct a

strange world of duality and unwillingly lead to the ideology of the vicious cycle
of poverty. (Oliveira, 2003, p. 31)

Such reproduction of "schemes learned in the Anglo-Saxon universities" (Oliveira, 2003, p. 32) would affect the whole way in which the Brazilian reality was to be apprehended, disseminating the conception of the country being an 'underdeveloped' nation (on similar critiques in other Latin American contexts see Sábato, 1975). Oliveira's criticism[8] is an example of the implications of the mere application of Eurocentric perspectives to realities alien to the West. Within the scope of Social Sciences and Humanities, a movement toward decentring the Western epistemological prevalence in its debates and toward a geopolitical pluralisation of theoretical bases grew steadily in the last half of the 20th century. Contributions contemplating the importance of the voices from the oppressed spread out among different parts of the globe and different fields (e.g. Freire, 1974; Hall, 1992; Herrera, 1971). The critique revolved around a particular addiction toward addressing social issues and marginalised groups from the lens of the canon of Western liberal democracies (Santos & Meneses, 2010). Those interpretations were often seen as generalising accounts, bringing everything under a global 'neoliberal factory' (Rexhepi, 2016).

More recently, during the early years of Big Data research, different groups of researchers alerted to the need for greater participation of the Global South in this new area of research (Arora, 2016). Most of all they claimed for theories of Big Data from the Global South[9]. One of the first attempts in trying to build such an agenda was the one-day conference Big Data from the South, held in Colombia in 2017. Its program brought several questions to light:

> How would datafication look like seen… 'upside down'? What questions would we ask? What concepts, theories, methods would we embrace or have to devise? What do we miss if we stick to the mainstream, Western perspective(s)? (Milan & Treré, 2017, p. 1)

Another major propeller of the initiative toward a Big Data from the South came more recently, in 2019, with the publication of a special issue on the theme in the journal Television & New Media. One of the first discussion topics in the "Big Data

8 Summarised briefly here under the risk of incurring in reductionism.
9 A movement born in the North for a Northern public, just for the record.

from the South" agenda revolved around what would be such things as North and South. In this sense, the initiative followed the theoretical foundations of decolonial studies (Mignolo & Escobar, 2010) by bringing the North/South divide to light as a way to search for "the absolute emancipation of all kinds of oppression and domination" involving power relations that privilege "legacies imposed by the colonial situation" and global capitalism. More than that, the initiative looked to foster "an innovative field of thought that privileges local epistemic elements" (Reis & Andrade, 2018, p. 3).

Despite recognising the importance of the geographical positioning of such situations of oppression, the "Big Data from the South" has avoided focussing their agenda on geographic boundaries. However, it is fundamental to notice the geographic space usually neglected - such space occupying most of the globe, as shown in Figure 3 and Figure 4.

> Firstly, there is the geographical South, i.e. the people, activities, politics, and technologies arising literally at the margins of the world as captured in the Mercator map. Secondly, and most importantly, our South is a place of (and a proxy for) resistance, subversion, and creativity. We can find countless Souths also in the Global North, as long as people resist injustice and fight for better life conditions against the impending 'data capitalism'. Our reflections on 'big data from the South' fit within—and hope to feed—the broader process of epistemological re-positioning of the Social Sciences. (Milan & Treré, 2017, p. 2)

However, more than having a critical instance regarding the current state of Big Data research, the "Big Data from the South" agenda has as its primary goal rereading the Big Data phenomena itself from the experiences, places, and theories of the South (Santos & Meneses, 2010). The project of rethinking Big Data from the South already has some interesting researches covering subjects such as political data movements (Chenou & Cepeda-Masmela, 2019), surveillance capitalism studies (Evangelista, 2017), and public policies in the South toward marginalised groups (Daly & Mann, 2018).
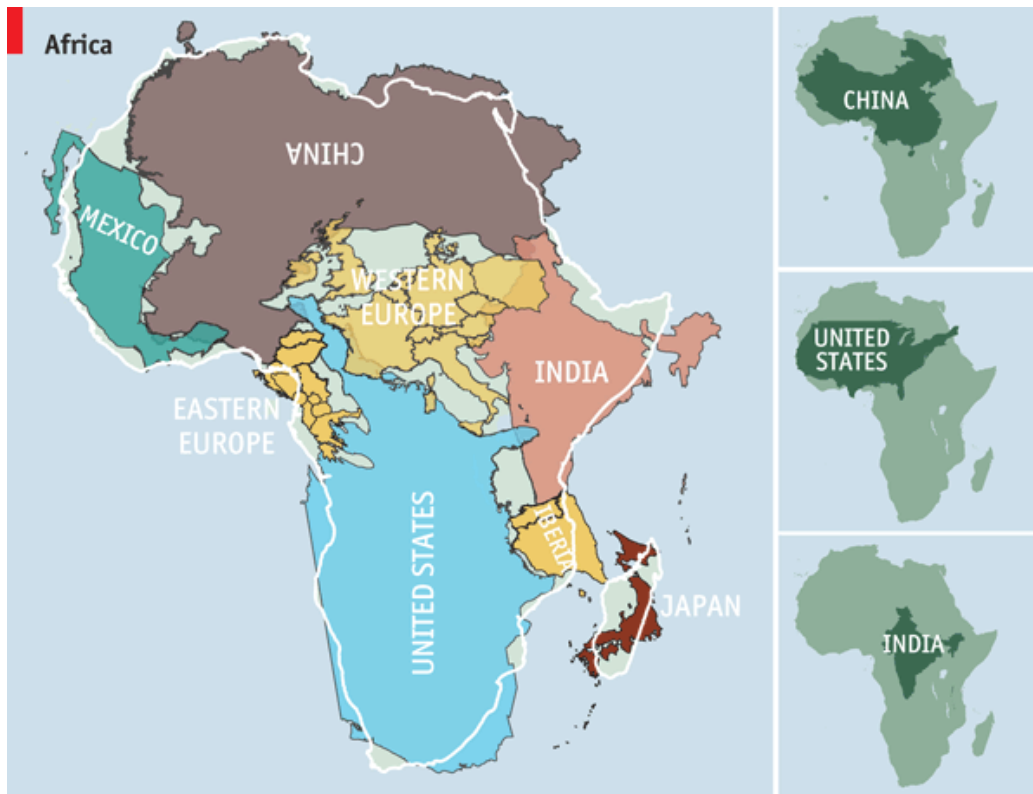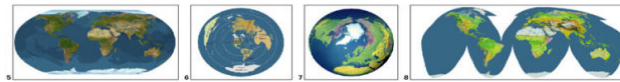
Figure 3. Map that criticises the exaggeration given in influential maps like the one of Gerardus Mercator, created in the context of the maritime explorations of European colonisers. Here, for example, the real proportion of the African continent compared to other parts of the globe. Source: "What's wrong with all our maps?". Available at https://bit.ly/tricolorbaiano

Figure 4. Map that criticises the exaggeration given in influential maps like the one of Gerardus Mercator, created in the context of the maritime explorations of European colonisers[10]. Source: "Big Data desde el Sur: El principio de una conversación que debemos tener". Available at https://data-activism.net/2017/10/desde-el-sur/

In short, the proposal to foster a Big Data from the South, although recent, finds an echo in flourishing literature proposing a critique of the Euroamerican epistemological supremacy in Big Data research in Social Sciences and building plural perspectives of Big Data. Aiming to contribute to the Big Data from the South project, this paper presents an analytical mapping of one of the journals that published the highest number of studies on the implications of Big Data to Social Sciences.

BIG DATA & SOCIETY: MAPPING AND DISCUSSION

One of the first findings of the study was the identification of a meagre presence of researchers affiliated to institutions outside North America and Europe among the Editorial Board, as shown in Figure 5. A fact that is seen in other Social Sciences' fields as well (e.g. Ganter & Ortega, 2019).

---

10 The map also plays with the notion of greatness commonly associated with the North (see Nelson & Simmons, 2009).

Geographical distribution
of Editorial Board members
Australia by affiliation
4   Asia
4

North America (US and Canada)
22

Europe
42

- North America (US and Canada)
- Europe
- Australia
- Asia

Figure 5. Geographical distribution of Editorial Board members by institutional affiliation, among a total of 72 people spread into editorial staff, supervisors and Board members. Of the total, 19 are affiliated to US institutions and 28 to British institutions. Updated information until October 2019[11].

The institutional affiliation of the authors who published in BD&S between 2014 and 2016 exposes a similar disparity, as we can see in Table 1. US institutions based a total of 45 of the 131 articles published by BD&S in the period, which accounts for more than a third of the total (34.35%). Along with British institutions, which account for a total of 39 articles (29.77%), the British-American slice reaches over 64% of the entire corpus, while only four papers come from researchers affiliated to institutions outside North America-Europe and Australia[12].

---

11 "Editorial Board". Available at https://us.sagepub.com/en-us/sam/journal/big-data-society
12 The unique situation of Australia and New Zealand in postcolonial discussions plus their historical position alongside traditional capitalist nations of the North can be seen in Mann & Daly (2018).

Table 1

Geographical distribution by institutional affiliation of the authors of the articles of the first six issues of BD&S. Source: Silva, 2019.

COUNTRIES	NUMBER OF PAPERS
US	45
United Kingdom	39
Netherlands	16
Canada	12
Germany	8
Ireland	5
Denmark	4
Australia	2
Belgium	2
France	2
Sweden	2
Austria	1
Croatia	1
Spain	1
Finland	1
Italy	1
Switzerland	1
United Arab Emirates	1
Singapore	1
Egypt	1
Japan	1

The analysis also covered the number of articles that took as their study object Latin American or African contexts. As shown in Figure 6 only three of the 131 articles published in the period (2.2%) take as their study object Southern contexts[13]: i) The paper from Mulder, Ferguson, Groenewegen, Boersma and Wolbers (2016), that addresses the way digital humanitarian groups used crowdsourcing

---

13 An exception is made here to the work of Cardullo (2015), which deals with the Twitter blockade made by the Turkish government in the face of demonstrations against Prime Minister Erdogan and the ways in which Twitter users have reacted to it. Cardullo (2015) built his analysis holding interviews with Turkish Twitter users, presenting specific characteristics of the way marginal groups aggrouped themselves in contexts of censorship on Twitter.

and open-source software during humanitarian crises in Nepal and Haiti; ii) Rieder, Abdulla, Poell, Woltering and Zack (2015) researching the Facebook page "We Are All Khalid", one of the main propellers of the political turmoil that led to the resignation of Hosni Mubarak in 2011; iii) And Nir Kshetri's (2014) paper on the benefits of 'successful' Big Data applications in 'developing' countries such as Kenya and Brazil.
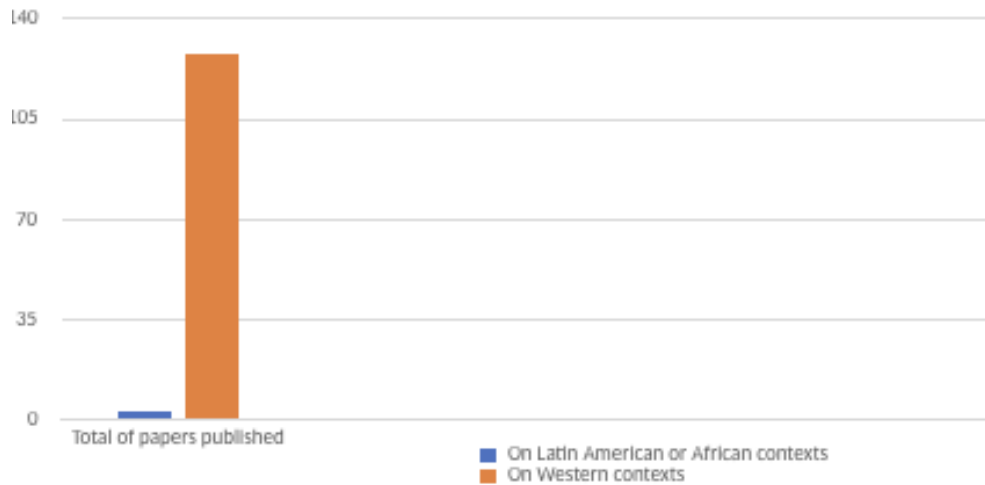

Figure 6. Distribution of papers by choice of study object/context concerning geopolitical affairs. Source: Silva, 2019.

Out of the three papers, however, only two cite local productions (Kshetri, 2014; Rieder et al., 2015). Mulder et al. (2016) mention Unicef reports, maps produced by US NGOs and even a Wired report but make no mention to initiatives, reports or papers from local agents. Rieder et al. (2015) include, among the co-authors, an Egyptian researcher who is cited in the material (Abdullah, 2014) together with another Egyptian author as well as anthropologists who did fieldworks in Syria. Among the three Kshetri (2014) is the one that most referrers to Southern Global actors, mentioning researches produced by people from Indonesia, India, Zimbabwe, and Zambia. His work argues for more extensive use of Big Data practices – 'consolidated' in Western multinationals - in 'developing' countries. None of the papers even mentions issues associated with the North/South divide and the asymmetrical power relations between the two.

The mapping identified a scenario of almost total invisibility of discussions about and from the South, that can be seen in the lack of authors from the South in the first three years of research on Big Data at BD&S. Such a conclusion is disturbing given the recognition that the period chosen for the analysis refers precisely to the initial discussions in the journal, a period of maturation of ideas and paths to be followed in the future. This scenario indicates that these paths have indeed departed from and directed themselves almost exclusively toward the intellectual islands of North America, Europe, and Australia/New Zealand – and their institutions.

Another troublesome conclusion is the fact that there are a low number of references to local literature in papers that deal with contexts of the Global South. That is precisely what Ganter & Ortega (2019, p. 79) refer to when they point to a "tendency to talk about rather than with" the South.

The Editorial Board of BD&S was contacted in order to inform the number of papers submitted by authors from Latin American or African institutions in the journal as well as geographic locations of the number of views and downloads made in the website. These numbers would show whether the absence of authors from the Global South in BD&S had any relation to disinterest from the part of Southern actors on the discussions or in the lack of knowledge of the journal itself. The answer came in the form of a survey on the total of submissions made to BD&S for the full run of the journal – up to July 2019. The data is for country assignments related to institutional affiliation (Table 2 below). With this setup, 16% of the submissions have been from scholars based in Southern institutions, and 8% of BD&S articles are from scholars based in Southern institutions.

Table 2

Geographical distribution by institutional affiliation of the submissions made to BD&S for the full run of the journal.

Country/Region	Accepted	Rejected	Total
Afghanistan	2	0	2
Argentina	0	1	1
Australia	16	10	26
Austria	3	2	5
Belgium	3	5	8
Canada	14	14	28
China	0	7	7
Croatia	1	0	1
Czech Republic	0	1	1
Denmark	14	8	22
Estonia	0	1	1
Finland	3	2	5
France	5	9	14
Germany	15	19	34
Ghana	0	1	1
Hong Kong	0	1	1
Hungary	0	2	2
India	1	17	18
Iran (the Islamic Republic of)	0	4	4
Iran, Islamic Republic of	0	1	1
Ireland	5	2	7
Israel	0	1	1
Italy	3	6	9
Japan	1	0	1
Kazakhstan	0	2	2
Kenya	0	1	1
Korea (the Republic of)	1	1	2
Lebanon	0	1	1
Malaysia	0	3	3

Country/Region	Accepted	Rejected	Total
Malta	0	1	1
Nepal	0	2	2
Netherlands	21	12	33
New Zealand	1	2	3
Nigeria	0	5	5
Norway	1	2	3
Oman	0	1	1
Philippines	0	1	1
Poland	0	2	2
Portugal	0	2	2
Russian Federation	0	1	1
Saudi Arabia	0	1	1
Singapore	0	1	1
Slovakia	0	1	1
South Africa	0	1	1
Spain	0	2	2
Sweden	7	5	12
Switzerland	2	9	11
Taiwan	1	0	1
Tunisia	0	1	1
Turkey	0	1	1
United Kingdom of Great Britain and Northern Ireland	56	41	97
United States	81	67	148

The editors were asked about possible reasons for such discrepancy, and answered that "a higher portion of articles we receive from scholars at Southern institutions (relative to scholars from Northern institutions) are desk rejected as they do not fit the scope of the journal". He went on to write that such papers avoided discussing social aspects of data[14]. It would be interesting to assess in-depth the reasons for such rejections, but no further answers were coming from the editors.

---

14 Message received by email on July 27, 2019.

With all that in mind, several factors need to compose the bigger picture for understanding the low participation of Southern Global authors, institutions and movements in BD&S publications, such as the language barrier. Despite the need for more research on this subject, as other researchers in different fields also stated (Ganter & Ortega, 2019), it is possible to conclude that the lack of Southern voices in BD&S is one of the main reasons for the low engagement of Big Data research in BD&S with the Global South.

GLOBAL SOUTH AND BD&S: LIMITS

To the findings obtained through the analytical mapping, it must also be taken into account the venues where BD&S resonate. One should not imagine, after all, that the papers published in BD&S correspond to the total picture of Big Data studies among social scientists. These papers circulate in specific locations and between specific communities, as seen in the previous topic. It must be clear primarily that all scientific research and publication processes are social (D'Andrea & Delich, 2005). The fact that the discussions of BD&S take place in the context of a Euroamerican environment is of vital importance to the scope of the points of view therein.

> Publications, especially articles in specialised journals, have become institutionalised as the 'ultimate' form of scientific communication. Specialised journals fulfil a key role in the scientific disciplines. They both secure the shared values of a scientific community and endorse what that community takes to be certified knowledge. (Vanderstraeten, 2010, p. 559)

Some theorists recognised that an author's entry into one of these academic environments and their publication media depends less on the quality of the research than on their ability to adapt to the 'language' and the canons of these communities (Vanderstraeten, 2010; Burt & Doreian, 1982). The same is especially true for 'periphery' academics, located outside the circulation venues of these publications and their communities, and subject to the woes of using a foreign language such as English for their endeavours (Dueñas, 2012).

Thus, the lack of Southern voices in BD&S, as well as the small number of works that address the Global North/South asymmetries, cannot be detached from the

place where BD&S is entrenched. A relevant question to ask, therefore, is where else are Big Data discussions and, more specifically, Big Data from the South circulating? For this question, however, the paper should offer no significant contributions, despite it being one of the further developments of this research (see Silva, 2019). Notwihstanding, it is essential to note that studies on Big Data from the South also find resonance in the Global South, apart from the European-led project outlined in the previous topic.

One of Latin America's leading technology and data research networks, Lavits (Latin American Network of Surveillance, Technology and Society Studies), has long been a hallmark for the study of matters such as surveillance capitalism and Big Data taking into account the historical relationship of asymmetry between the North and the Global South. The Network's last symposium, for example, had as its central theme "Asymmetries and (in)visibilities" reuniting social scientists from four different Latin American countries[15].

Although there is no journal in the Global South with the same explicit editorial purpose as BD&S, i.e. to address Big Data from a sociological standpoint, many kinds of research produced by Southern scholars take into account the North/South divide. Such works revolve around surveillance capitalism (Cruz, 2017; Firmino, Cardoso, & Evangelista, 2019) and personal data protection policies (Zanatta, 2017; Silveira, Avelino, & Souza, 2016). It is important to note, therefore, that the absence of Southern voices and the Global North/South divide at BD&S does not mean that it is not relevant elsewhere. Nevertheless, the diagnosis of such reality in a journal whose purpose revolves around understanding Big Data through critical sociology lenses points to a struggle of these communities of social scientists in providing insights beyond what their canon and situation offer (Silva, 2019).

CONCLUSION

The study, although limited in scope, brings questions that aim to foster future researches. The Global South invisibility in the articles of BD&S highlight implications of problems in Big Data research in Social Sciences that compromise the entire epistemic foundations of the field. In response to the concerns raised at the beginning

15 More details are available at Lavits website (Retrieved in December 10, 2019) https://lavits.org/eventos/simposio-lavits-2019/?lang=en

of the paper, the survey found a scenario of Euroamerican hegemony both in the number of institutions and authors involved in BD&S works and in the epistemological content. The lack of approaches coming from the Global South results in the tendency to talk about the South rather than through or with Southern actors. The paper wanted to point out the importance of decentralising Big Data studies, currently almost entirely tied to analysis from the Euroamerican canon.

Given such a situation, this research posits itself in the context of the "Big Data from the South" agenda. Its claim is for not only broader epistemic participation of the Global South in the discussions around Big Data, but also for other formulations of this phenomenon coming from other lenses, like those from the South with ideas born out of the margin. Further studies should consider which movements are already articulating Big Data from the South into the Global South, perhaps through an analytical mapping in the line of what accomplished here. After all, what is Big Data in Southern contexts? The Big Data experienced in the margins is the same as Laney's 3Vs [volume, speed, and variety] (2001) extended to many other Vs by Uprichard (2013)?

This paper argues for a change in the way the traditional questions around Big Data are built, claiming for an effort to understand data implications from marginal contexts (Milan & Treré, 2019). Moving beyond the traditional liberal accounts of Big Data might be the first step, as Mann, Devitt, & Daly (2019, p. 9) stated: "if digitisation and data are inevitabilities, then we have to (re)imagine the kind of digitised world and data we want to see rather than only offering a naysaying critique of the status quo".

REFERENCES

Akoka, J., Comyn-Wattiau, I., & Laoufi, N. (2017). Research on Big Data: A systematic mapping study. *Computer Standards & Interface, 54*(2), 105-115.

Arora, P. (2016). Bottom of the data pyramid: Big Data and the Global South. *International Journal of Communication*, *10*(2016), 1681-1699.

Boyd, D., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication and Society, 15*(5), 662-679.

Burt, R.S., & Doreian, P. (1982). Testing a structural model of perception: Conformity and deviance with respect to journal norms in elite sociological methodology. *Quality and Quantity, 16*(2), 109-150.

Chenou, J.M., & Cepeda-Másmela, C. (2019). #NiUnaMenos: Data activism from the Global South. *Television & New Media, 20*(4), 396-411.

Cruz, L.R. (2017). Google Suite for Education e o avanço do capitalismo de vigilância sobre as tecnologias educacionais. In P. Peña, R. Garrido, C.S. Baeza, R. Firmino, M.M. Kanashiro, & F. Bruno (Eds.), *Anais V Simpósio Internacional LAVITS: "Vigilância, Democracia e Privacidade na América Latina: vulnerabilidades e resistências"* (pp. 390-399). Retrieved from http://lavits.org/publicacoes/anais-2017/?lang=pt

D'Andrea, L., & Declich, A. (2005). The sociological nature of science communication. *Journal of Science Communication*, *4*(2), 1-9.

Dueñas, P.M. (2012). Getting research published internationally in English: An ethnographic account of a team of Finance Spanish scholar's struggles. *Ibérica*, *24*(1), 139-156.

Evangelista, R. (2017). Capitalismo de vigilância no Sul Global: Por uma perspectiva situada. In P. Peña, R. Garrido, C.S. Baeza, R. Firmino, M.M. Kanashiro, & F. Bruno (Eds.), *Anais V Simpósio Internacional LAVITS: "Vigilância, Democracia e Privacidade na América Latina: vulnerabilidades e resistências"* (pp. 243-253). Retrieved from http://lavits.org/publicacoes/anais-2017/?lang=pt

Firmino, R.J., Cardoso, B.V., & Evangelista, R. (2019). Hyperconnectivity and (Im)mobility: Uber and surveillance capitalism by the Global South. *Surveillance & Society, 17*(1/2), 205-212.

Freire, P. (1974). *Pedagogia do oprimido*. São Paulo, SP: Paz e Terra.

Ganter, S.A., & Ortega, F. (2019). The Invisibility of Latin American scholarship in European Media and Communication Studies: Challenges and opportunities of de-westernization and academic cosmopolitanism. *International Journal of Communication*, *13*(2019), 68-91.

Hall, S. (1992). The West and the rest: Discourse and power. In S. Hall & B. Gieben (Eds.), *Formations of modernity* (pp. 275-331). London, United Kingdom: Polity Press.

Herrera, A. (1971). *Ciencia y política en América Latina*. Ciudad de México, Mexico: Siglo XXI.

Iliadis, A., & Russo, F. (2016). Critical Data Studies: An introduction. *Big Data & Society*, 3(2), 1-7.

Kaplan, F. (2015). A map for big data research in digital humanities. *Frontiers in Digital Humanities, 2*(1), 1-7.

Kshetri, N. (2014). The emerging role of Big Data in key development issues: Opportunities, challenges, and concerns. *Big Data & Society, 1*(2), 1-20.

Kukutai, T., & Taylor, J. (Eds.) (2016). *Indigenous Data Sovereignty: Toward an agenda*. Canberra, Australia: ANU Press.

Laney, D. (2001, February 6). *3D data management: Controlling data volume, velocity, and variety*. Retrieved from http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf

Mann, M., & Daly, A. (2018). (Big) Data and the North-in-South: Australia's informational imperialism and digital colonialism. *Television & New Media*, *20*(4), 379-395.

Mann, M., Devitt, S.K., & Daly, A. (2019). Introduction. In A. Daly, S.K. Devitt & M. Mann (Eds.), *Good Data* (pp. 8-24). Amsterdam, Netherlands: Institute of Network Cultures.

Mignolo, W.D., & Escobar, A. (Eds.). (2010). *Globalization and the decolonial option*. London, United Kingdom; New York, NY: Routledge.

Milan, S., & Treré, E. (2019). Big Data from the South(s): Beyond data universalism. *Television & New Media, 20*(4), 319-335.

Milan, S., & Treré, E. (2017, October 16). Big Data from the South: The beginning of a conversation we must have. Retrieved from http://bit.ly/ST2ALavits2019

Mulder, F., Ferguson, J., Groenewegen, P., Boersma, K., & Wolbers, J. (2016). Questioning Big Data: Crowdsourcing crisis data towards an inclusive humanitarian response. *Big Data & Society, 3*(2), 1-13.

Nelson, L.D., & Simmons, J.P. (2009). On Southbound Ease and Northbound Fees: Literal Consequences of the Metaphoric Link between Vertical Position and Cardinal Direction. *Journal of Marketing Research, 46*(6), 715-724.

Oliveira, F. (2003). *Crítica à Razão Dualista – O Ornitorrinco*. São Paulo, SP: Boitempo Editorial.

Reis, M.N., & Andrade, M.F.F. (2018). O pensamento decolonial: Análise, desafios e perspectivas. *Revista Espaço Acadêmico, 17*(202), 1-11.

Rexhepi, P. (2016). Liberal luxury: Decentering Snowden, surveillance and privilege. *Big Data & Society, 3*(2), 1-3.

Rieder, B., Abdulla, R., Poell, T., Woltering, R., & Zack, L. (2015). Data critique and analytical opportunities for very large Facebook Pages: Lessons learned from exploring "We are all Khaled Said". *Big Data & Society, 2*(2), 1-22.

Sábato, J.A. (Ed.). (1975). *El pensamiento latinoamericano en la problemática ciencia-tecnología-desarrollo-dependencia*. Buenos Aires, Argentina: Paidós.

Santos, B.S., & Meneses, M.P. (Eds.). (2010). *Epistemologias do Sul*. São Paulo, SP: Editora Cortez.

Silva, G.C. (2019). A questão do Big Data nas Ciências Sociais: Panorama inicial a partir da Big Data & Society. *Anais do Seta*, *9*(1), 43-57.

Silveira, S.A., Avelino, R., & Souza, J. (2016). A privacidade e o mercado de dados pessoais. *Liinc em Revista, 12*(2), 217-230.

Tomaz, T., & Silva, G.C. (2018). Repensando big data, algoritmos e comunicação: Para uma crítica da neutralidade instrumental. *Revista Parágrafo, 6*(1), 31-42.

Uprichard, E. (2013, October 1). *Big data, little questions. Discover Society*. Retrieved from http://discoversociety.org/2013/10/01/focus-big-data-little-questions/

Vanderstraeten, R. (2010). Scientific communication: Sociology journals and publication practices. *Sociology, 44*(3), 559-576.

Wienhofen, L.W.M., Mathisen, B.M., & Roman, D. (2015, September 10). *Empirical Big Data research: A systematic literature mapping*. Retrieved from https://arxiv.org/pdf/1509.03045.pdf

Ylijoki, O., & Porras, J. (2016). Conceptualizing Big Data: Analysis of case studies. *Intelligent Systems in Accounting, Finance and Management, 23*(4), 295-310.

Youtie, J., Porter, A.L., & Huang, Y. (2017). Early social science research about Big Data. *Science and Public Policy, 44*(1), 65-74.

Zanatta, R.A.F. (2017). Proteção de dados pessoais como regulação de risco: Uma nova moldura teórica? In C.B. Israel, D.J. Vicentin, F.R. Rosa, J.C.F. Santos, N.S. Patrício, R. Gatto, & R. Segurado (Eds.), *Anais REDE 2017: I Encontro da Rede de Pesquisa em Governança da Internet* (pp. 175-193). Retrieved from http://redegovernanca.net.br/public/conferences/1/anais/Anais_REDE_2017-1.pdf