Geography in the big data age: an overview of the historical resonance of current debates

Daniela Ferreira & Mário Vale

GEOGRAPHY IN THE BIG DATA AGE: AN OVERVIEW OF THE HISTORICAL RESONANCE OF CURRENT DEBATES

Title: Geography in the big data age: an overview of the historical resonance of current debates

Authors: Daniela Ferreira*; Mário Vale

Affiliation: Centre of Geographical Studies, Institute of Geography and Spatial Planning, Universidade de Lisboa, Lisbon, 1600-276.

E-mail: danielaferreira2@campus.ul.pt; mario.vale@campus.ul.pt

ABSTRACT. Although much has been written about the challenges of big data, there has been little reflection on the historicity of such debates and what we can learn from it. With this in mind, the aim of this article is to situate the epistemological debates over big data in geography historically. We focus on the three most relevant topics in current discussions around big data that have significant historical resonance, namely its methodological challenges, its scientific value, and its positionality. We conclude by arguing that understanding the historical resonance of current big data debates is helpful to find new ways to question its epistemological consequences. *Keywords: big*

data, digital geography, geographical epistemology, GIS, history of geography.

The term big data has been used to designate or explain the existence of the most varied forms of new technologies for storage, collection, and data management, which are characterized by its high speed, variety, and volume, as well as for its flexibility, comprehensiveness, resolution, and indexing (Kitchin 2014a). The emergence of big data generated a turn in geographical epistemology, raising new discussions on epistemological challenges and opportunities (Kitchin 2013, 2014a; Kitchin and Dodge 2014). On the one hand, it has been argued that big data allows a new form of doing science based solely on data processing, often referred to as data-driven science, which has the capacity of explaining everything through data analysis (Kelling and others 2009; Miller 2010; Kitchin 2014b). Others have argued that big data has revived positivism, which can now be the dominant epistemological paradigm (Wyly 2014; Kitchin 2014b; Thatcher and others 2018). On the other hand, several concerns have been raised over the emergence of big data, namely regarding the methodological challenges it poses, the scientific value of big data processing, and the neutrality and positionality of the data it produces.

These debates are not entirely new in geography. Indeed, they echo longstanding polemics regarding the place of the digital in the discipline. The digital has been a part of geographic practice for a long time and has assumed different roles as it was appropriated by different epistemological perspectives over the history of the discipline. Initially, the digital supported the production of knowledge as a tool for the quantitative methods of spatial science (Tobler

1959; Balchin and Coleman 1967; Tomlinson 1968; Goodchild 2018). These digital tools were gradually embraced by geographers working with geographic information systems (GIS) who proclaimed GIS as a science during the 1990s (Mark 2003; Fisher 2007; Goodchild 2004a, 2010). However, the use of the digital in geography has also been a target of several critiques by critical geographers during the last decades, most notoriously by feminist geographers (Elwood 2000, 2006; Leszczynski 2009). Meanwhile, as the proliferation of digital spaces and practices that marked the information age took place, geographers became increasingly interested in its impact on society (Bell 1973, 1979; Hall and Preston 1988; Hepworth 1990; Kellermann 1993; Goddard 1992, 1995; Li 1995; Graham and Marvin 1996; Castells 2000, 2004).

Although much has been written about the challenges of big data and how it has been understood differently by distinct epistemological perspectives (Ash and others 2016; Thatcher and others 2018), there has been little reflection on the historicity of such debates and what we can learn from it. This has been pointed out by Trevor Barnes (2013), who argues that history deserves more attention in the big data epistemological debates because big data is a consequence of relevant historical steps such as the quantitative revolution. From this viewpoint, big data is not a completely new paradigm, but rather a continuation (Bell and others 2009; Leszczynski 2014).

This article aims to situate the epistemological debates over big data in geography historically. We focus on the three most relevant topics in current discussions around big data that have significant historical resonance. Firstly, we approach the methodological challenges of big data, and compare them to

the challenges of training and acquisition of new methods that geographers faced during the dominance of spatial science. For many, big data is a challenge for training and methods learning because current statistical methods are unfit to analyze the volume of information that is constantly being updated (Kitchin 2013; Thatcher and others 2018). Interestingly, geographers training during the spatial science paradigm felt similar difficulties to assimilate the necessary methods for data processing (Barnes 2001, 2003). Secondly, we relate the debates on the scientific value of big data processing with the vibrant debates on GIS as a science (Pickles 1997; Mark 2003; Goodchild 2004a; Wright and others 2004; Fisher 2007). Currently, big data is challenging the concept of science as some argue that it is possible to produce knowledge and explain phenomena solely through data processing methods (Barnes 2013; Kitchin 2013, 2014b). Similarly, the discipline of GIS has engaged in debates about whether the use of digital tools has to be subject to scientific methodologies, or if a data-driven geographic practice is possible (Schuurman and Pratt 2002; Barnes 2001, 2003; Ash and others 2016). Thirdly, we approach the critiques to big data geography and relate them to the critiques on the positionality of GIS. Recently, it has been argued that the outputs produced through big data methods are not neutral and relate to its situated perspective. These critiques echo the discussions about positionality and situated knowledge undertaken by critical geographers, which were often aimed at the practice of GIS (Bondi and Domosh 1992; Smith 1992; Keller and Longino 1996; Dixon and Jones 1998; Schuurman and Pratt 2002; Kwan 2002; Leszczynski 2009).

We will show that there are important historical resonances in the geographic debates about big data. In this sense, we argue that there are consistent epistemological issues that reemerge over time, which provides novel insights on the scope of these debates.

This article is further divided into five sections. First, the emergence of big data as a challenge to geography's epistemology is approached. The second section compares the challenges that geographers faced during the dominance of the spatial science paradigm to become proficient in quantitative methods with the challenges of training and methodology that big data has generated. The third section relates the discussions about the need for scientific theories in the big data age with the debates around GIS as a science. The fourth section looks into the continuity of the critiques to the objectivity of spatial science, GIS, and big data in geography. We conclude by arguing that understanding the historical resonance of current big data debates is helpful to find new ways to question its epistemological consequences.

<<A-Head>> DATA REVOLUTION AND THE EMERGENCE OF THE BIG DATA AGE

Geographers became increasingly concerned with the impact of digital technologies in society during the 1990s. A large number of works on the geographies of information and technology were published in this decade, including studies on the geography of information and regional and urban development (Goddard 1995; Gillespie and Williams 1988), the geographies of computers and information technologies (Hepworth 1990; Li 1995), the geographies of telecommunications (Kellermann 1993; Graham and Marvin 1996), and the geography of innovation (Feldman 1994).

More recently, geographers have shifted their attention toward the data revolution, such as the rapid development of the capacity to collect, store, and process digital data (Graham and Shelton 2013; González-Bailón 2013). The data revolution is characterized by the automatic mode through which data are captured and produced, made possible by the development and the proliferation of several technological innovations, infrastructures, and techniques. According to Barnes (2013), we have gone from an "avalanche of numbers," as coined by Ian Hacking (1990), to an "avalanche of big data." Other authors consider that we have entered the big data age (Kitchin 2013, 2014a; Kitchin and Dodge 2014) or the petabyte age (Anderson 2008). Although there is no specific definition for big data, some authors have identified its main features (Boyd and Crawford 2012; Zikopoulos and others 2012; Mayer-Schonberger and Cukier 2013). Big data consist in data sets characterized by a huge volume of information that is exhaustive in scope and has great resolution in detail. These data sets are flexible in the edition or creation of their fields and relational, as it is possible to conjugate different data sets. Big data have also been more briefly defined by the three Vs: high volume, high velocity, and high variety (Kitchin 2013; Mayer-Schonberger and Cukier 2013; Miller and Goodchild 2015). In this context, data are now ubiquitous, as nearly every social, economic, political, and environmental process is being digitized and stored by sensors and software. The production of huge amounts of data has been generated predominantly by actors such as app developers, retail chains, financial institutions, mobile phone operators, security firms, and internet companies. These actors are usually associated with large corporations, and it has been noted that the data access, storage, and

collection process is generally privatized (Thatcher and others 2018). Most often, big data sets are commercialized to firms. In this sense, they constitute business opportunities that allow the expansion of markets and the construction of new marketing strategies.

Big data has become important to develop new modes of analysis and knowledge production (Graham and Shelton 2013). In recent years, there has been an increase of works on this topic. Places are now often understood through the dense and complex layers of representation that are created, accessed, and filtered through codified algorithms (Thrift and French 2002; Goodchild and others 2007; Graham 2010; Kitchin and Dodge 2014; Wilson 2017).

Although this transition from the information age to the big data age is widely acknowledged, the epistemological consequences of this data surge for geography are still a matter of contention. Big data play an important role in epistemological discussions in geography because a great deal of spatial information has been produced and this offers a significant potential for studies in this field.

<<A-Head>> THE DIGITAL CHALLENGES TO GEOGRAPHY'S TRAINING AND METHODS

Big data has brought new methodological challenges for geographers as it requires specific techniques of data processing that are not traditionally part of a geographer's training. This situation resembles the challenges that geographers faced during the dominance of the spatial science paradigm to become proficient in quantitative methods. Quantification in geography emerged as an alternative to the descriptive approaches of regional geography

that dominated the discipline during most of the first half of the twentieth century (Capel 1983). The inspiration for this approach stemmed from the German works of regional economics (Webber 1928; Christaller 1966; von Thünen 1966). However, geography's turn to quantification was fuelled by the need to conceive the discipline as a true science. A significant debate on the methodology of geography took place in the United States, where authors such as Fred Schaefer (1953) argued that geographers needed to apply scientific methodologies if they wished to be considered a true science (Cresswell 2013). Quantification was seen as the way to provide geography with the scientific theories it needed. At the time, the Lund School in Sweden and the pioneering work of Torsten Hägerstrand (1967) became an important reference as a geographical approach that used quantification and statistical modelling (Barnes 2001). Their research approach became popular in the Anglophone world, and geographers began creating spatial models and applying statistical mathematics to find spatial patterns and cause-effect relations.

During this period, the digital emerged in geography as a tool, as computers were intensively used to undertake new forms of statistical analysis and spatial modelling (Haggett 1965; Hägerstrand 1967). Several statistical methods required massive amounts of calculation, and so computers became a valuable tool to process the models of spatial science. Moreover, computer software and coding also provided possibilities to develop new techniques related to mapping and spatial modelling. Thus, computers were fundamental for those "seeking ways efficiently to measure, store, describe, and display data concerned with geographic surfaces, conditions, and processes" (Pruitt

1979, 107). Barnes (2001) highlights the arrival of the large Friden tabletop calculator and the IBM computer as important tools for the development of quantitative geography. Other nondigital technologies were also important for geographers at this point, namely the machines for reproducing text such as the duplication machine or the mimeograph (Barnes, 2001).

The introduction of these technologies in the practice of geography required preparing geographers to master them as part of their training. While some geographers transitioned from working under the paradigm of regional geography to applying mathematical methods of the physical sciences and neoclassic economics in their work, some of the most notable names of spatial science emerged from geography departments that invested in advanced statistical courses, such as the case of the University of Washington (Barnes 2001, 2003). On the other hand, this period is also marked by a generation that came to geography with degrees in hard sciences, such as mathematics and physics. Grasping techniques of computer programming was an important part of this investment, although it was often necessary to "bootstrap" this learning as training was not easily available (Barnes 2001, 2003). It must be noted that while some were willing to incorporate new quantitative epistemologies despite not having the technical skills, others simply refused to accept these epistemologies as valid in the first place.

The impact of the computer in geography extended beyond the capability of large-scale calculation, as it molded the discipline's language itself. Barnes (2001) called our attention to the way that the vocabulary of regional geography was substituted by a new vocabulary characterized by terms of computer programming such as "if, then," numerical data and

operations of correlations and regressions. Likewise, João Rua and Maurício Abreu (2004) noted that the language of quantitative geographers during the 1960s was different from the language of other geographers as it became increasingly filled with computer and statistical terms.

The use of computers in geography would eventually lead to the first project of digital mapping and the first national cartography initiatives (Tobler 1959; Balchin and Coleman 1967). Simultaneously, structures of spatial information were built, such as the Canada Geographic Information System, and since the 1960s the development of GIS took place (Tomlinson 1968; Goodchild 2018). Therefore, digital technologies were used as a tool for underpinning the development of positivist spatial sciences. While the digital allowed geographers to develop projects with social and political relevance, the tool also had a profound impact in the discipline itself, in the sense that it shaped the training and language of geographers.

Geographers face the need to dominate new data processing methods if they wish to grasp the potential of the sources of big data. As in the period of quantitative revolution, geographers are coming to the realization that they need new methods to analyze these new data sets because traditional statistical methods are not enough (Kitchin 2013; Thatcher and others 2018). This implies new training for geographers. A set of skills related to data processing are necessary, and research centers are being created in order to prepare new geographers. The need to learn about big data methods is recognized across the social sciences, not only due to its usefulness for future research, but also because it is a popular topic that attracts potential funding (Sui and Delyser 2012; DeLyser and Sui 2013a, 2013b). As Rob Kitchin argues, "[t]he

challenge for human geography is to push back against such naïve forms of predatory science and, at the same time, to be able to compete in the same channels for research funding" (2013, 264).

Geographers are once again learning and adapting to the challenges imposed by epistemological shifts toward quantification. While big data analytics, where advanced analytic techniques operate on big data sets, Philip Russom (2011, 5) have emerged as group of methods that geographers linked to GIS have embraced to support the production of knowledge, others have shown that there is no unique way to use big data and that learning quantitative data processing methods is not the only pathway to use big data in geography. It has been shown that it is possible to use big data sources away from algorithms and without applying standard big data analytics. For instance, some studies in the field of cultural geography and the geohumanities have approached big data focusing on the cocreation of digital public art content through the online engagement of users (Kidd 2014; Zebracki and Luger 2018). Tess Osborne and others (2018) applied alternative methods that combine spatialized big data and art moves into qualitative narratives and representations of space.

There are several approaches to big data that provide alternative ways to analyze its content. While there are emerging methodological challenges due to the emergence of big data, geographers have been successful at readjusting traditional methods to these new data sources (Zebracki and Luger 2018).

The historical resonance of the methodological and training challenges of big data must lead us to question its long-lasting consequences for the

discipline. As some geographers turn to big data analytics or engage with more creative approaches to big data sources, it is important to understand how this turn might be changing not only the training of geographers, but also the language and vocabulary of geography, as did the digital quantification of spatial science.

<<A-Head>> BETWEEN SCIENCE AND TOOL

Big data is generating debates about the epistemological shifts it may cause in geography and other disciplines. These debates have similarities with the GIS\GISc (geographic information science) discussions that took place during the 1990s. In both cases, digital techniques that were seen as tools to support scientific methods eventually led to discussions about the need for scientific theories. Right now, it is unclear what kind of science big data can support, which has generated a myriad of hypothesis about its epistemological position. In a similar manner, the use of GIS also sparked discussion about whether GIS should be understood as a mere tool for data processing or as a scientific endeavor based in positivism.

This discussion began in the early 1990s, as Michael Goodchild (1992) argued that GIS should be seen not only as a tool for science, but as a science itself. Traditionally, GIS was seen as a way of solving problems, rather than a scientific approach that it is able to understand and discover problems. The process of claiming GIS as a science was made step-by-step and we can observe the change in its designations over time: from geographic information system to science of geographic information systems, and later science of geographic information (Pickles 1997). Moreover, some changes in journals showed how GIS was being rethought in academic research. For example,

International Journal of Geographical Information Systems changed its name in 1997 to *International Journal of Geographical Information Science* (Fisher 2007). It is also important to mention that the National Center of Geographic Information and Analysis (NCGIA) coined the term of "geography of the information society" in 1995, and some years later changed it to "science of geographic information."

According to John Pickles (1997), a theoretical turn was taking place in GIS, in which issues were raised about the rationality of GIS and the open, contested nature of science. Several authors contributed to this discussion regarding what sort of theory and epistemological framework should underpin GIS (Goodchild 2004a, 2010). Goodchild (2004a) argued that GIS was a true science underpinned by Anselin's concept of spatial heterogeneity, which stated that every phenomenon is unevenly distributed across space, and Tobler's first law of geography, which highlighted the role of proximity in the relations between spatial elements (Tobler 1970; Goodchild 2004a). Other proposals seemed to put a greater emphasis on the technical and theoretical knowledge that allows the development of GIS as a research tool (Reitsma 2013). Goodchild himself described GISc as "the storehouse of knowledge that is implemented in GIS and that makes GIS possible" (2004a, 710), and Caron and others (2008) defined it as the foundation upon which GIS technologies are built.

A significant debate was also raised about the relationship of GISc with other disciplines. Some believed GISc could be understood as a subset of information science, and that we could find its components in a myriad of scientific fields such as statistics or computer science (Goodchild and others

1999; Goodchild 2004b). However, Jonathan Raper (2009) argued that GISc exceeds this definition, and could be considered a multidiscipline with a consistent theory, data, and software engineering work.

Similarly to GIS, big data entered geography mainly as a new tool to gather and analyze data, but also led to discussions about the need for scientific theories. For instance, Luciano Floridi (2012) and Sabina Leonelli (2012) have argued that we must not only think about big data as a data deluge and a technique, but also rethink its philosophy of science. These authors have pointed out the emergence of new modes of positivism and empiricism under the context of the big data age. It has been argued that scientific theories are no longer needed as big data makes data-driven science possible. The again, some argue that big data has renovated quantitative positivism.

The belief that big data means the end of theory or the revival of empiricism stems from the idea that it is no longer necessary to create new models or theories, or to look for causalities in spatial phenomena. Instead, it is argued that big data offers all the necessary data to explain and understand spatial phenomena, and that it is possible to find patterns and construct knowledge through correlations between these data (Gorman 2013; Kitchin 2013, 2014b). Danah Boyd and Kate Crawford (2012) argue that there is a mythology regarding big data, in that they provide novel forms of analysis and, as such, become a new and superior form of intelligence that makes theories and models obsolete. In this scenario, knowledge would be "born from the data rather than born from the theory" (Kitchin 2014b, 6). Authors such as Theodore Porter (1996), Chris Anderson (2008), Marc Prensky (2009), and Bruno Strasser (2012) have argued that big data have the capacity to speak

for themselves and do not need contextual or a domain-specific knowledge to make sense of the analysis or interpretations. Similarly, geographers have also engaged with the emergence of a quantitative empiricism, in which all knowledge is produced through the capture, storage, and analysis of data. According to this perspective, we are entering a new era of computational social science or a kind of "data-driven" science, in which potential questions and hypotheses would be derived from knowledge discovery techniques (Kitchin 2014b). In this way, data-driven science would provide additional, valuable insights that traditional science does not generate (Miller 2010).

Now, a new quantitative positivism is emerging, underpinned by the idea that positivism gained a new momentum with these new methods of data processing. For Elvin Wyly (2014), we are entering an automatized positivism, in which the emergence of big data presents an opportunity for positivist scholars to access more integrated information (González-Bailón 2013; Kitchin 2014b; Thatcher and others 2018).

In sum, like GIS, big data is also raising questions about the need for theory. However, the history of GISc shows that despite the apparent autonomy that some techniques and methods can achieve in terms of producing research results, theoretical discussions remain relevant to establish foundational knowledge that allows the development of geography. In this sense, the reappearance of positivism seems to be one of the epistemological consequences of big data, but there have been critical approaches to big data in geography that must be highlighted.

<<A-Head>> Critical Geographies

As big data provides new quantitative techniques and generates a newfound interest in positivist approaches, critiques about the objectivity and neutrality of big data also emerge. These critiques strike a chord with past concerns regarding spatial science and the development of GIS.

While the debate about what would constitute a science of geographical information was sprouting in the 1990s, a significant critique to spatial science and GIS was also taking place by the hands of critical geographers, who argued that geography was "retreating from knowledge to information" (Taylor 1990, 212). Although such debates overlap (Barnes 2009), these later critiques must be highlighted as a predecessor of contemporary critiques to big data.

An important starting point for this discussion was John Harley's (1989) critical cartography and the idea of deconstruction of the map. Harley (1989) argued that maps should be distinguished from territory itself, as they are both a representation of territory and a technology that contributes toward normalizing and legitimizing power. The idea that the map does not convey neutral knowledge and has political consequences was likewise applied to GIS. Another significant contribution was Donna Haraway's (1988) critique of the "God Trick," which refers to the supposed neutrality of abstract scientific understandings of the world. Haraway (1988) argued that there is no such thing as a purely neutral, detached, and universal perspective, and that knowledge always comes from a situated perspective. These critiques to mapping and science were used to criticize the claims for the neutrality and objectivity of GIS. For instance, Lake argued that "[p]ositivist assumptions of

objectivity, value-neutrality and the ontological separation of subject and object (or of the analyst and the object of analysis) constitute epistemological conditions with political and ethical consequences" (1993, 405). This debate produced a high number of publications about GIS, and its role on geography and epistemology (Schuurman 1999, 2000). In addition, the term cartographic anxiety was created by Derek Gregory (1994) as a critique to objectivism in mapping (Painter 2008). Cartographic anxiety emerges from the gaps between governmental and citizen representations, which create an anxiety related to the fact that the spatial imaginaries do not correspond to cartographic representation (Painter 2008). In other cases, this anxiety might come from ethnic, religious, or political minorities (Cons 2016), but also might be related with nostalgic mourning for past territories (Cartier 2013) or the existence of poetic and corporealized visions of the nation-state (Ramaswamy 2010).

The major critique came by the hand of feminist geographers and it was related to the relationship between the State and the use of GIS. Many critiques were aimed at the colonialist militarism, masculine positivism, and the cartographic rationalities of GIScience, and it was argued that this form of knowledge was ethnocentric (Bondi and Domosh 1992; Dixon and Jones 1998; Schuurman and Pratt 2002; Chrisman 2005; Leszczynski and Elwood 2014). Moreover, most critics believed that GIS served the needs and interests of largest corporations and a Western masculinist influence had prevailed while smaller groups were excluded (Taylor 1991; Smith 1992; Lake 1993; Pickles 1995, 1997).

It was argued that GIS needed a more consistent reflection on its practices and the consequences of its technologies on society (Schuurmaan

and Pratt 2002). However, critics of the claims for the neutrality and objectivity of GIS did not aim to exclude these practices from geography. Instead, they argued that a more reflexive and representative approach to the GIS was necessary (Smith 1992; Lake 1993; Pickles 1993). While authors such as Goodchild (2004a, 2010) remained faithful to the positivist approach, the notion that GIS needed to be more inclusive was well received across the field.

Critical and qualitative GIS established itself as a significant subfield as these debates took place during the 1990s (Schuurman 2006). Critical cartography contributed to the flourishing of critical GIS, based on the critiques by feminist geographers, who challenged the neutrality of GIS (Smith 1992; Keller and Longino 1996; Kwan 2002; Leszczynski 2009). While feminist scholars pointed out that technologies were being used to support ideologies that were considered censurable, such as colonialism, critical GIS demonstrated that GIS can also be used to produce positioned, reflective, nonmasculinist, emotional/affective, and inclusive knowledge (Elwood 2000; Knaw 2002; Schuurman 2002, 2006; Pavlovskaya 2006). Critical and qualitative GIS allowed subaltern geographies to become more visible, by contributing to increase the capacity of vulnerable social groups to construct geographic knowledge.

Mordechai Haklay and others (2008) and Haklay (2013) argued that all the biases of positivist geography also apply to big data. In fact, some of the critiques made to spatial science and GIS by critical geographers during the 1980s and 1990s are now being directed to big data. For instance, the issue of inclusion\exclusion in big data has been raised, and it has been noted that

those that cannot be counted (if they are not online, for instance) will not appear on databases. As in the period of spatial science, concerns about neutrality and objectivity have been highlighted. The previously unreachable knowledge acquired through the analysis of data sets with high volume, velocity, and variety has been questioned from the point of view of its quality, neutrality, and veracity. The production of big data through volunteered geographic information has sparked epistemological discussions among academics about its quality, veracity, and neutrality (Elwood and others 2012; James Ash and others 2016; Paulo Cambra and others 2019)---discussions that are similar to those raised during the 1990s about GIS. Although big data sets have been conceived as a neutral, rational, and objective representation of reality, some authors have been focusing their attention on the biases of data sets. For instance, it has been argued that the production of knowledge can be seized by a specific group, namely the giants of digital capitalism that are defining who produces knowledge and for whom (Kitchin 2014a).

Another critique has been aimed at the notion that numbers can speak for themselves through assumptions that they embody (Anderson 2008; Prensky 2009; Croll 2012). It has been argued that correlations can respond to the "what," but not to the "why" (Crawford 2013). Moreover, there is a risk of understanding the world only indirectly and through numbers---a critique that was also aimed at positivism once (Gould 1981; Haraway 1991; Amin and Thrift 2002). Mark Graham and Taylor Shelton (2013) argued that it is impossible for the data to speak for themselves because data are inevitably associated not only to assumptions, but also to the experiences, values, and ideas that cannot be captured by numbers. This critique was likewise directed

to the quantitative revolution during the late 1970s and early 1980s (Sayer 1984).

Graham and Shelton (2013) have argued that big data can have negative consequences in social sciences if it creates a new scientistic, positivist, and quantitative turn. Although big data should be recognized as a new contribution for science, it is important not to forget the potential of qualitative and critical data studies.

The new momentum of positivism and its principles is generating critiques related to the biased and positioned knowledge of big data. It is important, however, to understand how these critiques can lead to a critical big data geography that is capable of opening new paths for participation in geographic studies, like the critiques to GIS lead to the emergence of critical GIS. There have been some first steps to achieve this. The use of volunteered geographic information in e-governance processes has contributed toward the inclusion of minorities in local governance, for instance. Despite such efforts, big data are producing digital divides as a consequence of the lack of participation, accentuating existing inequalities. According to Anran Yang and others (2016), information inequalities on OpenStreetMap have increased in countries with low data output. It has been argued that the use of technologies is predominantly associated with social segments that have higher incomes, higher education levels, and social and economic capital (Benjamin 2001).

<<A-Head>> CONCLUSION

This article sought to situate the epistemological debates over big data in geography historically, establishing bridges between current discussions on the epistemology of big data in geography and the challenges that spatial

science created for the training of geographers, the debates on the value of theory in GIS, and the critiques made by several epistemological approaches to spatial science and GIS (Barnes 2013; Kitchin 2013, 2014b). This historical approach helps us to understand the depth of the debates around big data, as it shows us how they echo longstanding polemics regarding the place of the digital in the discipline. More importantly, understanding the historical resonance of current big data debates is also important to find new ways to question its consequences. This study has led us to question the consequences of big data in the language and vocabulary of geography, the status of theory in the discipline, and the emergence of a critical big data geography.

Although the epistemological issues raised by big data debates have historical resonance, it is important to highlight that the contexts are different. It is especially noteworthy that big data emerges in a period in which several epistemological approaches coexist in geography, as for this reason the debates around the epistemology of big data are far less contingent in scope. While in the past the digital was mainly present in spatial science or in the field of GIS, big data has had an impact on the various fields of geography. With this, big data seems to have revived several longstanding concerns about geographic theory and provoked a response by distinct approaches. Big data also presents an opportunity to bridge quantitative and qualitative approaches in physical geography, human geography, or GIS.

Big data is a digital mark that can be able to bridge the different fields in geography, creating and combining novel forms of work and generating new opportunities of research. Geographers are already actively creating novel forms of producing knowledge with big data, even without applying big data

analytics (Kidd 2014; Rose and Willis 2018; Osborne and others 2018). These efforts to combine and adjust traditional methods to new data sources, bridging qualitative and quantitative approaches (Zebracki and Lugar 2018), must be recognized so that geography can make use of its epistemological diversity to develop different visions about big data and its epistemological consequences.

The appearance of big data is more than a new digital development. It had a significant impact on the epistemology of geography, as it had the capacity to generate new concepts. The broad character of big data and its transversality in the different fields of geography has led geographers to question the current paradigms in geography and these discussions were able to move different epistemological perspectives in geography. Big data has generated new epistemological turns related to both quantitative and critical geography. It is important to conclude that the digital has never had a greater epistemological impact on geography than big data. This must lead us to acknowledge the need to further understand the depth of the uncertain relations between geographic epistemologies and digital technologies.

<<A-HEAD>> REFERENCES

Amin, A., and N. Thrift. 2002. *Cities: Reimaging the Urban*. London: Polity.

Anderson, C. 2008. *The End of Theory: The Data Deluge Makes the Scientific Method Obsolete.*

http://www.wired.com/science/discoveries/magazine/16-07/pb_theory.

Ash, J., R. Kitchin, and A. Leszczynski. 2016. Digital Turn, Digital Geographies? *Progress in Human Geography* 1--19.

Balchin, W., and A. Coleman. 1967. Cartography and Computers. *The Cartographer* 4: 120--127.

Barnes, T. J. 2001. Lives Lived and Lives Told: Biographies of Geography's Quantitative Revolution. *Environment and Planning D: Society and Space* 19 (4): 409--429.

_____. 2003. The Place of Locational Analysis: A Selective and Interpretive History. *Progress in Human Geography* 27 (1): 69--95.

Barnes, T. 2007. "Not Only … But Also": Quantitative and Critical Geography. *The Professional Geographer* 61 (3): 292--300.

_____. 2013. Big Data, Little History. *Dialogues in Human Geography* 3 (3): 297--302.

Bell, D. 1973. *The Coming of Post-Industrial Society: A Venture in Social Forecasting*. Harmondsworth, U.K.: Penguin.

_____. 1979. The Social Framework of the Information Society. In *The Computer Age: A Twenty-Year View,* edited by M. Dertouzous and L. Moses, 163--211. Cambridge, Mass.: MIT Press.

Bell, G., T. Hey, and A. Szalay. 2009. Beyond the Data Deluge. *Science* 323 (5919): 1297--1298.

Benjamin, M. 2001. *The Digital Divide: Facing a Crisis or Creating a Myth?* Cambridge, Mass.: MIT Press.

Bondi, L., and M. Domosh. 1992. Other Figures in Other Places: On Feminism, Postmodernism, and Geography. *Environment and Planning D: Society and Space* 10 (2): 199--213.

Boyd, D. and K. Crawford. 2012. Critical questions for big data. *Information. Communication and Society* 15 (5): 662--679.

Cambra, P. J., A. Gonçalves, and F. Moura. 2019. The Digital Pedestrian Network in Complex Urban Contexts: A Primer Discussion on Typological Specifications. *Finisterra-Revista Portuguesa de Geografia* LIV (110): 155--170.

Capel, H. 1983. *Filosofía y Ciencia en la Geografía Contemporánea.* Contemporánea. Barcelona: Barcanova.

Caron, C., D. Goyer, S. Roche, and A. Jaton. 2008. GISc Journals Ranking and Evaluation: An International Delphi Study. *Transactions in GIS* 12: 293--321.

Cartier, C. 2013. What's Territorial about China? From Geopolitical Narratives to the 'Administrative Area Economy.' *Eurasian Geography and Economics* 54 (1): 57--77.

Castells, M. 2000. *The Rise of the Network Society, The Information Age: Economy, Society and Culture*. Vol. I. Oxford: Blackwell Publishers.

Castells, M. 2004. *A Galáxia Internet, Reflexões sobre a Internet, Negócios e Sociedade.* Lisboa: Fundação Calouste Gulbenkian.

Chrisman, N. 2005. Full Circle: More than Just Social Implications of GIS. *Cartographica: The International Journal for Geographic Information and Geovisualization* 40 (4): 23--35.

Christaller, W. 1966. *Central Places in Southern Germany*. London: Prentice Hall.

Cons, J. 2016. *Sensitive Space: Fragmented Territory at the India-Bangladesh Border*. Seattle: University of Washington Press.

Crawford, K. 2013. *The Hidden Biases of Big Data*. Harvard Business Review Blog, 1. https://hbr.org/2013/04/the-hidden-biases-in-big-data.

Cresswell, T. 2013. *Geographic Thought. A Critical Introduction*. Oxford: Wiley-Blackwell.

Croll, A. 2012. Big Data is Our Generation's Civil Rights Issue, and We Don't Know It. In *Big Data Now: Current Perspectives from O'Reilly Radar,* edited by O'Reilly Radar Team, 55--59. Sebastopol, Calif: O'Reilly Media.

DeLyser, D., and D. Sui. 2013a. Crossing the Qualitative Quantitative Divide II: Inventive Approaches to Big Data, Mobile Methods, and Rhythmanalysis. *Progress in Human Geography* 37 (2): 293--305.

_____. 2013b. Crossing the Qualitative Quantitative Chasm III: Enduring Methods, Open Geography, Participatory Research, and the Fourth Paradigm. *Progress in Human Geography* 38 (2): 294--307.

Dixon, D. P., and J. P. Jones. 1998. My Dinner with Derrida, or Spatial Analysis and Poststructuralism Do Lunch. *Environment and Planning A* 30 (2): 247--260.

Elwood, S. 2000. *Information for Change: The Social and Political Impacts of Geographic Information Technologies*. Minneapolis: University of Minnesota Press.

_____. 2006. Beyond Cooptation or Resistance: Urban Spatial Politics, Community Organizations, and GISbased Spatial Narratives. *Annals of the Association of American Geographers* 96 (2): 323--341.

Elwood, S., M. Goodchild, and D. Sui. 2012. Researching Volunteered Geographic Information: Spatial Data, Geographic Research, and New Social Practice. *Annals of the Association of American Geographers* 102 (3): 571--590.

Feldman, M. P. 1994. *The Geography of Innovation*. Dordrecht, The Netherlands: Kluwer.

Fisher, P. 2007. Valediction. *International Journal of Geographical Information Science* 21: 1165--1170.

Floridi, L. 2012. Big Data and Their Epistemological Challenge. *Philosophy and Technology* 25 (4): 435--437.

Gillespie, A., and H. Williams. 1988. Telecommunications and the Reconstruction of Regional Comparative Advantage. *Environment and Planning A* 20 (10): 1311--1321.

Goodchild, M. F. 1992. Geographical Information Science. *International Journal of Geographic Information Systems* 6: 31--45.

_____. 2004a. GISc, Geography, Form, and Process. *Annals of the Association of American Geographers* 94 (4): 709--714.

_____. 2004b. The Validity and Usefulness of Laws in Geographic Information Science and Geography. *Annals of the Association of American Geographers* 94 (2): 300--303.

_____. 2010. Twenty Years of Progress: GISc in 2010. *Journal of Spatial Information Science* 1: 3--20.

_____. 2018. Reimagining the History of GIS. *Annals of GIS* 24 (1): 1--8.

Goodchild, M. F., M. Egenhofer, K. Kemp, D. Mark, and E. Sheppard. 1999. Introduction to the Varenius Project. *International Journal of Geographic Information Science* 13: 731--745.

Goodchild, M. F., M. Yuan, and T. J. Cova. 2007. Towards a General Theory of Geographic Representation in GIS. *International Journal of Geographical Information Science* 21: 239--260.

Goddard, J. 1995. *ICTs Space and Place: Theoretical and Policy Challenges*. Workshop on Informatics and Telecom Tectonics: Information Technology, Policy, Telecommunications and the Meaning of Space.

_____. 1992. New Technology and the Geography of the UK Information Economy. In *Understanding information Business,*

Technology and Geography, edited by K. Robbins, 178-201. London: Belhaven.

González-Bailón, S. 2013. Big Data and the Fabric of Human Geography. *Dialogues in Human Geography* 3 (3): 292--296.

Gorman, S. 2013. The Danger of a Big Data Episteme and the Need to Evolve Geographic Information Systems. *Dialogues in Human Geography* 3 (3): 285--291.

Gould, P. 1981. Letting the Data Speak for Themselves. *Annals of the Association of American Geographers* 71 (2): 166--176.

Graham, M. 2010. Neogeography and the Palimpsests of Place. *Tijdschrift voor Economische en Sociale Geografie* 101 (4): 422--436.

_____. 2013. Geography/Internet: Ethereal Alternate Dimensions of Cyberspace or Grounded Augmented Realities? *The Geographical Journal* 179 (2): 177--182.

Graham, M., and T. Shelton. 2013. Geography and the Future of Big Data, Big Data and the Future of Geography. *Dialogues in Human Geography* 3 (3): 255--261.

Graham, S. 1998. The End of Geography or the Explosion of Place? Conceptualizing Space, Place and Information Technology. *Progress in Human Geography* 22 (2): 165--185.

Graham, S., and S. Marvin. 1996. *Telecommunications and the City: Electronic Spaces, Urban Places*. London: Routledge.

Gregory, D. 1994. *Geographical Imaginations*. Oxford: Blackwell.

Hacking, I. 1990. *The Taming of Chance*. Cambridge, U.K.: Cambridge
University Press.

Hägerstrand, T. 1967. *Innovation Diffusion as a Spatial Process*. Chicago:
University of Chicago Press.

Haggett, P. 1965. *Locational Analysis in Human Geography*. London: Edward
Arnold.

Haklay, M. 2013. Neogeography and the Delusion of Democratization.
Environment and Planning A 45 (1): 55--69.

Haklay, M., A. Singleton, and C. Parker. 2008. Web Mapping 2.0: The
Neogeography of the GeoWeb. *Geography Compass* 2 (6): 2011--
2039.

Hall, P., and P. Preston. 1988. *The Carrier Wave: New Information
Technology and the Geography of Innovation, 1846--2003*. London:
Unwin Hyman.

Haraway, D. 1988. Situated Knowledges: The Science Question in Feminism
and the Privilege of Partial Perspective. *Feminist Studies* 14 (3): 575--
599.

_____. 1991. *Simians, Cyborgs and Women: The Reinvention of Nature*.
New York: Routledge.

Harley, J. B. 1989. Deconstructing the Map. *Cartographica* 26 (2): 1--20.

Hepworth, M. 1990. *Geography of the Information Economy*. New York: Guilford.

Keller, E. F., and H. E. Longino. 1996. *Feminism and Science.* Oxford: Oxford University Press.

Kellermann, A. 1993. *Telecommunications and Geography*. London: Belhaven.

Kidd, J. 2014. *Museums in the New Mediascape: Transmedia, Participation, Ethics*. Farnham, U.K.: Ashgate.

Kitchin, R. 2013. Big Data and Human Geography: Opportunities, Challenges and Risks. *Dialogues in Human Geography* 3 (3): 262--267.

_____. 2014a. *The Data Revolution. Big Data, Open Data, Data Infrastructures and Their Consequences*. London: Sage.

_____. 2014b. Big Data, New Epistemologies and Paradigm Shifts. *Big Data and Society* 1 (1): 1--12.

Kitchin, R., and M. Dodge. 2014. *Code/Space. Software and Everyday Life*. Cambridge, Mass.: MIT Press.

Kwan, M. P. 2002. Feminist Visualization: Re-envisioning GIS as a Method in Feminist Geographic Research. *Annals of the Association of American Geographers* 92 (4): 645--661.

Lake, R. W. 1993. Planning and Applied Geography: Positivism, Ethics, and Geographic Information Systems. *Progress in Human Geography* 17: 404--413.

Leonelli, S. 2012. Introduction: Making Sense of Data-driven Research in the Biological and Biomedical Sciences. *Studies in History and Philosophy of Biological and Biomedical Sciences* 43 (1): 1--3.

Leszczynski, A. 2009. Rematerializing GIScience. *Environment and Planning D: Society and Space* 27: 609--615.

_____. 2014. On the Neo in Neogeography. *Annals of the Association of American Geographers* 104 (1): 60--79.

Leszczynski, A., and S. Elwood. 2014. Feminist Geographies of New Spatial Media. *Canadian Geographer* 59: 12--28.

Mark, D. M. 2003. *Geographic Information Science: Defining the Field.* London: Taylor and Francis.

Mayer-Schonberger, V., and K. Cukier. 2013. *Big Data: A Revolution that Will Transform How We Live, Work and Think*. London: John Murray.

Miller, H. J. 2010. The Data Avalanche Is Here. Shouldn't We Be Digging? *Journal of Regional Science* 50 (1): 181--201.

Miller, H. J., and M. F. Goodchild. 2015. Data-driven Geography. *GeoJournal* 80 (4): 449--461.

Osborne, T., E. Warner, P. I. Jones, and B. Resch. 2018. Performing Social Media: Artistic Approaches to Analysing Big Data. *GeoHumanities*: 1--13.

Painter, J. 2008. Cartographic Anxiety and the Search for Regionality. *Environment and Planning A* 40: 342--361.

Pavlovskaya, M. 2006. Theorizing with GIS: A Tool for Critical Geographies? *Environment and Planning A* 38 (11): 2003--2020.

Pickles, J. 1995. *Ground Truth: The Social Implications of Geographic Information Systems*. New York: Guilford.

_____. 1997. Tool or Science? GIS, Technoscience and the Theoretical Turn. *Annals of the Association of American Geographers* 87: 363--72.

Porter, T. 1996. *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life.* Princeton, N.J.: Princeton University Press.

Prensky, M. 2009. H. Sapiens Digital: From Digital Immigrants and Digital Natives to Digital Wisdom. *Innovate* 5 (3): Article 1.

Pruitt, E. L. 1979. The Office of Naval Research and Geography. *Annals of the Association of American Geographers* 69: 103--108.

Ramaswamy, S. 2010. *The Goddess and the Nation: Mapping Mother India*. Durham, N.C.: Duke University Press.

Raper, J. 2009. *Annual Review of Information Science and Technology.* Medford, N.J.: Information Today.

Reitsma, F. 2013. Revisiting the 'Is GIScience a science?' Debate (or Quite Possibly Scientific Gerrymandering). *International Journal of Geographical Information Science* 27 (2): 211--221.

Rose, G., and A. Willis. 2018. Seeing the Smart City on Twitter: Colour and the Affective Territories of Becoming Smart. *Environment and Planning D: Society and Space:* 1--17.

Rua, J., and M. Abreu. 2004. Construindo uma boa relação com a História. *Geo UERJ* 15--16: 193--216.

Russom, P. 2011. *Big Data Analytics*. Towi Research.

Sayer, A. 1984. *Method in Social Research: A Realist Approach*. London: Hutchinson.

Schaefer, F. 1953. Exceptionalism in Geography: A Methodological Examination. *Annals of the Association American Geographers* 43 (3): 226--249.

Schuurman, N. 1999. Critical GIS: Theorizing an Emerging Discipline. *Cartographica* 36 (4): 1-99.

_____. 2000. Trouble in the Heartland: GIS and Its Critics in the 1990s. *Progress in Human Geography* 24 (4): 569--290.

_____. 2002. Women and Technology in Geography: A Cyborg Manifesto for GIS. *The Canadian Geographer* 46 (3): 258--265.

_____. 2006. Formalization Matters: Critical GIS and Ontology Research. *Annals of the Association of American Geographers* 96 (4): 726--739.

Schuurman, N., and G. Pratt. 2002. Care of the Subject: Feminism and
   Critiques of GIS. *Gender, Place and Culture: A Journal of Feminist
   Geography* 9 (3): 291--299.

Smith, N. 1992. History and Philosophy of Geography: Real Wars, Theory
   Wars. *Progress in Human Geography* 16: 257--71.

Strasser, B. J. 2012. Data-driven Sciences: From Wonder Cabinets to
   Electronic Databases. *Studies in History and Philosophy of Biological
   and Biomedical Sciences* 43 (1): 85--87.

Sui, D., and D. DeLyser. 2012. Crossing the Qualitative-Quantitative Chasm I:
   Hybrid Geographies, the Spatial Turn, and Volunteered Geographic
   Information (VGI). *Progress in Human Geography* 36 (1): 111--124.

Taylor, P. J. 1990. GIS. *Political Geography Quarterly* 9: 211--212.

_____. 1991. Further Thoughts on Geography and GIS. *Environment
   and Planning A* 23: 1087--1090.

Thatcher, J., J. Eckert, and A. Shears. 2018. *Thinking Big Data in Geography.
   New Regimes, New Research.* Lincoln: University of Nebraska Press.

Thrift, N., and S. French. 2002. The Automatic Production of Space.
   Transactions of the Institute of British Geographers 27 (3): 309--335.

Tobler, W. R. 1959. Automation and Cartography. *The Geographical Review*
   49 (4): 526--534.

_____. 1970. A Computer Movie Simulating Urban Growth in the
   Detroit Region. *Economic Geography* 46: 234--240.

Tomlinson, R. F. 1968. A Geographic Information System for Regional

Planning. In *Land Evaluation,* edited by G. A. Stewart, 200--210.

Melbourne, Australia: Macmillan.

von Thünen, J. 1826. [1966]. *Der Isolierte Staat in Beziehung auf*

Landwirtschaftslehre und Nationalökonomie, Translated by C. M.

Wartenberg, [Von Thünen's Isolated State]. Oxford: Pergamon Press.

Wilson, M. 2017. *New Lines: Critical GIS and the Trouble of the Map*.

Minneapolis: University of Minnesota Press.

Wright, D., M. Goodchild, and J. Proctor. 2004. Demystifying the Persistent

Ambiguity of GIS as 'Tool' Versus 'Science'. *Annals of the*

Association of American Geographers 87 (2): 346--362.

Wyly, E. 2014. Automated (Post)Positivism. *Urban Geography* 35 (5): 669--

690.

Yang, A., H. Fan, N. Jing, Y. Sun, and A. Zipf. 2016. Temporal Analysis on

Contribution Inequality in OpenStreetMap: A Comparative Study for

Four Countries. ISPRS *International Journal of Geo-Information* 5 (1):

5.

Zebracki, M., and J. Luger. 2018. Digital Geographies of Public Art: New

Global Politics. *Progress in Human Geography* 1--20.

Zikopoulos, P. C., C. Eaton, D. deRoos, T. Deutsch, and G. Lapis. 2012.

Understanding Big Data. New York: McGraw Hill.