

## A QUESTÃO DO BIG DATA NAS CIÊNCIAS SOCIAIS: PANORAMA INICIAL A PARTIR DA BIG DATA & SOCIETY

Guilherme Cavalcante Silva<sup>1</sup>

**Resumo:** Considerando o crescente nível de interesse para com o que se convencionou chamar de Big Data, que perpassa desde a pesquisa acadêmica até políticas governamentais e estratégias de acumulação de capital, este artigo apresenta os apontamentos iniciais de um panorama sobre os estudos em Big Data a partir de perspectivas dos diversos campos das Ciências Sociais. Devido ao tamanho e diversidade dos estudos nestas disciplinas, o trabalho toma como escopo de análise as seis primeiras edições da revista Big Data & Society, pioneira no fomento da investigação do Big Data entre os cientistas sociais. O objetivo é compreender, ao menos de forma inicial, como o Big Data é articulado nos variados campos das Ciências Sociais. Isto é, entender como se formam as tendências, lugares de fala e cortes temáticos em torno do tema, fazendo isso a partir do recorte da Big Data & Society. O artigo inicialmente introduz os debates em torno do Big Data, através de uma revisão da literatura recente acerca do assunto. Em seguida, apresenta distintos lugares de circulação do Big Data, com destaque para a produção científica. Por fim, explicita brevemente os motivos de escolha do objeto de pesquisa, o processo metodológico do levantamento e os detalhes iniciais do mapeamento analítico feito na revista. A pesquisa se utilizou de uma estratégia metodológica de análise quali-quantitativa. O artigo se encerra situando caminhos da pesquisa em Big Data nas Ciências Sociais e oferecendo questionamentos na tentativa de potencializar possibilidades de investigação na interface Big Data-Ciências Sociais.

**Palavras-chave:** Big Data. Ciências Sociais. Big Data & Society.

**Abstract:** Considering the evergrowing interest towards what has been usually called Big Data that permeates from the academic researches up to government policies and marketing strategies, this paper presents the initial notes from an overview on Big Data studies within the different perspectives of the Social Sciences field. Due to the size and diversity of this particular field, the work takes as its investigation scope the first six issues of the scientific publication Big Data & Society, precursor on the promotion of Big Data research among social scientists. Comprehending initially, at least, how Big Data is articulated in the diverse Social Sciences fields is the main goal of this paper. In other words, it seeks to understand how the places of speech, trends and thematic worries around Big Data are molded in the Social Sciences from the analysis of Big Data & Society. The paper initially addresses the concept of Big Data, synthesizing the debates around it over the last few years through a literature review. Then it will briefly explicit the main reasons for choosing Big Data & Society as the study object of this work, addressing the methodological processes and the initial details of the analytical mapping done in the publication as well. This research made use of a quali-quantitative analysis strategy. The paper then finishes situating research paths on the Big Data research in Social Sciences, offering questions on the attempt to potentialize investigations in the Big Data-Social Sciences interface.

**Keywords:** Big Data. Social Sciences. Big Data & Society.

---

<sup>1</sup> Mestrando em Divulgação Científica e Cultural no Laboratório de Estudos Avançados em Jornalismo (Labjor), Unicamp, e-mail: [guilhermecavalcantesilva@outlook.com](mailto:guilhermecavalcantesilva@outlook.com). Bolsista CAPES.

## Introdução

Exposições sobre o que seria o Big Data costumemente têm girado em torno de termos como estes, especialmente ao redor das palavras que Laney (2001) chamou de três Vs: velocidade, variedade e volume<sup>2</sup>. Na verdade, faltaria espaço para compartilhar as *v-words* já utilizadas para dar conta do que seria o Big Data: versatilidade, volatilidade, virtuosidade, vitalidade, vigor, virilidade, valor, veracidade, variabilidade, entre tantas outras (KITCHIN; MCARDLE, 2016). Todas aparecem em contraste com uma situação anterior, na qual as informações obtidas sobre consumidores, hábitos de vida e o uso do tempo por parte da população poderiam ser obtidos apenas em pequenas amostras agrupadas em pequenas bases de dados. Todavia, é importante lembrar que chamar os dados de grandes, complexos ou variados ainda diz pouco sobre o que são os dados. Como afirmam Boyd e Crawford (2012, 663): “Big Data é, de diversos modos, um termo pobre”.

A partir de incômodos proporcionados pelas problemáticas do Big Data na atualidade, e em busca de construções outras de Big Data que não as elencadas acima, este artigo oferece um panorama inicial de um levantamento sobre as pesquisas em Big Data, considerando-o a partir de perspectivas de pesquisadoras(es) dos diversos campos das Ciências Sociais. O objetivo é compreender quais são as tendências, como se formam os fundamentos e lugares de fala das tratativas sobre Big Data nesta área do saber em especial.

Considerando a enormidade da tarefa, o artigo se concentra num levantamento em torno do material publicado entre os anos de 2014 e 2016 na revista *Big Data & Society* (BD&S), publicação pioneira na exploração do *Big Data* a partir das contribuições das Ciências Sociais. O levantamento categorizou aspectos dos 131 artigos publicados no período investigado pela análise, como: autores(as), título dos artigos, palavras-chave, país de origem da pesquisa (da instituição a que se filiam os autores), áreas de pesquisa dos(as) autores(as), objetos de análise, principais referências utilizadas nos artigos, e escolhas metodológicas.

O artigo inicialmente introduz os distintos lugares de circulação do Big Data, com destaque para a produção científica. Após isso, explicita brevemente os motivos de escolha do objeto de pesquisa, os fundamentos metodológicos do levantamento e os detalhes iniciais do mapeamento analítico feito na revista. Por fim, o trabalho se encerra situando caminhos da pesquisa em Big Data nas Ciências Sociais e oferecendo questionamentos na tentativa de

---

<sup>2</sup> Um exemplo disso é o documento oficial do projeto recente da Google chamado Data Transfer Project e que engloba também gigantes da internet como Facebook, Twitter e Microsoft, com o propósito de reunir sob uma única plataforma toda a base de dados de usuários dos serviços das quatro companhias. No arquivo oficial deste que já é o maior projeto de portabilidade de dados do qual se tem notícia, não deixam de aparecer termos como “velocidade”, “eficiência”, “tamanho” e “transparência” como resultados naturais de tal iniciativa. Disponível em: < <https://datatransferproject.dev/dtp-overview.pdf> > Acesso em: 30/01/2019.

potencializar possibilidades de investigação na interface Big Data-Ciências Sociais. É importante ressaltar que o trabalho a seguir constitui uma versão inicial da pesquisa de investigação de Mestrado do autor.

### **1. Os campos do Big Data**

Apesar de recentes, os estudos em Big Data têm mobilizado um amplo espectro de esforços disciplinares e transdisciplinares. Um mapeamento sistemático realizado por pesquisadores de instituições superiores francesas no ano de 2016 tratou de investigar o terreno por onde circularam as pesquisas em Big Data entre os anos de 2013 e 2016 dentro da base de dados do ScienceDirect (AKOKA et al., 2017). Chama a atenção, além do número de artigos levantados, um total de 1843, a amplitude dos estudos em Big Data – com 24 áreas distintas representadas - bem como a presença marcante da Ciência da Computação e das Engenharias entre as publicações, alcançando 935 artigos dentre o total (50,73%), mais do que metade de todas as outras 22 disciplinas somadas. Isso ajuda a explicar o fato de aspectos técnicos e performativos [qualidade, eficiência, velocidade etc] aparecerem entre os termos mais pesquisados, ainda segundo Akoka et al. (2017).

Além dos diversos locais disciplinares e transdisciplinares a se ter em conta na pesquisa em Big Data, é preciso reconhecer também a importância de um quadro geopolítico acerca da produção científica sobre o Big Data. Apesar da tradição acadêmica de países europeus como Inglaterra, Alemanha e França, por exemplo, bem como os Estados Unidos, é importante ressaltar a importância de movimentos como o da China, que ao menos na área de Finanças e Administração tem se equiparado aos Estados Unidos em produção acadêmica em Big Data - e não apenas no âmbito acadêmico, vale lembrar.

A partir da Figura 1, percebe-se, desde já, uma forte tendência de concentração euroamericana, bem como a expressividade do Big Data em locais como China, Índia, Coreia do Sul e Japão. Por outro lado, o Sul global, em especial a América Latina e África, aparecem em um nível de quase inexpressividade na balança do Big Data. Esta desigualdade impacta conceitos, teorias e métodos que aparecem no horizonte do que é o Big Data atualmente (MILAN; TRERÉ, 2017).

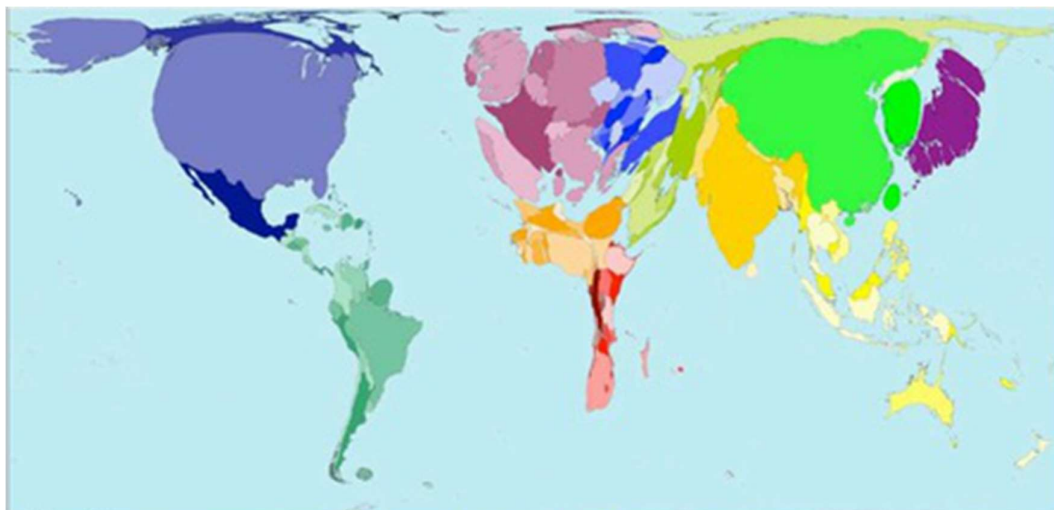


Figura 1: Mapa geopolítico da balança global do Big Data. Elaborado pela cientista computacional Stéphane Grumbach, do Institut National de Recherche en Informatique et en Automatique (Inria/França)<sup>3</sup>

As principais constatações de tal mapeamento são a: i) interdisciplinaridade do campo de estudos do Big Data, se é que pode-se falar de um ‘único’ campo ou mesmo da existência de um ‘campo’ do Big Data; e, mesmo com a amplitude dos esforços em torno do tema, ii) a concentração teórica nos campos técnicos da Ciência da Computação e das Engenharias, bem como iii) a centralização euroamericana e chinesa na produção, uso e concepção do Big Data. Este é o alcance da mobilização em torno do Big Data na contemporaneidade. Todavia, o olhar neste artigo se volta para uma visão mais específica do que é e de como circula o Big Data: a das Ciências Sociais.

## 2. O objeto de estudo

O relacionamento entre Big Data e Ciências Sociais tem sido avaliado com mais intensidade na literatura científica recente. Nos últimos dois anos, por exemplo, duas obras foram publicadas tendo como temática o “uso de Big Data nas várias disciplinas das Ciências Sociais” (CHEN; YUN, 2018, p. 3). Como os próprios nomes dos trabalhos indicam<sup>4</sup>, a perspectiva é a do uso de metodologias quantitativas baseadas em extração e análise de dados para as pesquisas empíricas por parte de cientistas sociais. O foco aí é em fornecer um apanhado geral dos principais métodos e, a partir daí, traçar um passo-a-passo para uma ciência social quantificável (FAZEKAS, 2014; FOSTER et al., 2017).

<sup>3</sup> “Big Data? The Global Imbalance”. Disponível em: <<http://www.in2p3.fr/actions/formation/Info13/lift-27sept-121001065613-phpapp01.pdf>> Acesso em: 30/01/2019.

<sup>4</sup> “Big data and social science: a practical guide to methods and tools”, de Foster et al. (2017); e “Big Data in Computational Social Science and Humanities”, de Chen (2018).

Todavia, o que fazer de um campo que sempre teve por objeto de análise algo tão complexo e irreduzível como este que se convencionou chamar de humano e um *habitat* tão subjetivo quanto fluido como o social? Que dizer de aspectos qualitativos, epistemológicos, filosóficos, políticos? Big Data é, para os cientistas sociais, apenas uma ferramenta analítica quantitativa (numérica), como fazem crer Foster et al. (2017)? Dados são apenas poderes neutros a serem domados por aqueles que detém capital técnico, material e financeiro? São questionamentos como estes que Dalton e Thatcher (2014) acreditam ser a grande contribuição que as Ciências Humanas e Sociais podem trazer para os estudos em Big Data, um engajamento crítico para com os dados, para além de sua mera aplicação funcional (SILVA, 2018).

De todo modo, independentemente da perspectiva adotada pelos cientistas sociais, o objetivo do levantamento foi perpassar a multiplicidade de abordagens para situar os caminhos da pesquisa em Big Data nas Ciências Sociais. Todavia, em um campo tão largo quanto este, o recorte escolhido para análise foi o de uma revista científica específica, visto que ela se constitui como um dos principais canais de divulgação, interação, desenvolvimento e consolidação dos estudos sobre Big Data (CONFORTO et al., 2011). Tal escolha se justifica ante ao ensejo da pesquisa de voltar seu olhar especificamente para a produção acadêmico-científica das Ciências Sociais em relação com o Big Data.

A escolha do objeto recaiu sobre a revista Big Data & Society (BD&S). A BD&S teve sua primeira edição publicada em abril de 2014 com a proposta de ser uma publicação pioneira na exploração do *Big Data* a partir das contribuições das Ciências Sociais<sup>5</sup>. Seu objetivo era “analisar as práticas de Big Data, ou envolver propostas empíricas [...] que também reflitam nas consequências de como as sociedades [com/através dos *big data*] são representadas [epistemologias], compreendidas [ontologias] e governadas [políticas]”, como aponta o editorial<sup>6</sup>. A proposta é perceptível também pela presença de importantes nomes das Ciências

---

<sup>5</sup> É importante ressaltar a existência de outras publicações científicas que se propõem, como mote principal, a investigar o *Big Data*. Em levantamento breve, constatei a presença de doze revistas acadêmicas, incluindo *Big Data & Society*, a lidarem, como proposta editorial, com o Big Data. São elas [entre colchetes, seus principais enfoques]: *IEEE Transactions on Big Data* [Engenharia e Ciência da Computação], *Big Data* [Pesquisa e inovações em big data - interdisciplinar], *Journal of Big Data* [Ciência da Computação e Computação cognitiva], *Big Data Research* [Interdisciplinar - com destaque para publicações na área de Ciência da Computação], *Big Data & Cognitive Computing* [Computação cognitiva], *Frontiers in Big Data* [Data mining e cybersecurity], *International Journal of Big Data Intelligence* [Vigilância e gerenciamento de *big data*], *Big Data Analytics* [Computação cognitiva], *International Journal of Big Data and Analytics in Healthcare* [Pesquisa em Saúde], *Open Journal of Big Data* [Progresso e inovação no/com o uso dos *big data* - interdisciplinar], e *Big Data Analytics for Healthcare* [Pesquisa em Saúde]. Várias outras publicações que tratam de temáticas relacionadas e/ou similares, como o *International Journal of Data Science and Analytics*, *Computational Statistics & Data Analysis (CSDA)*, *Data Science Journal*, e *Journal of Digital Humanities*, são citadas aqui, porém permanecem fora do escopo do trabalho. A escolha pela análise da Big Data & Society se deve pela singularidade de sua proposta editorial voltada às Ciências Sociais, que é o lócus da pesquisa.

<sup>6</sup> About the Journal. Big Data & Society. Disponível em: <<http://bigdatasoc.blogspot.com.br/p/big-data-and-society.html>> Acesso em: 30/01/2019.

Sociais e Humanidades Digitais entre os membros do comitê consultivo da revista, tais como Rob Kitchin, Bruno Latour, Tarleton Gillespie e Luciano Floridi.

Até dezembro de 2018, 213 artigos foram publicados pela Big Data & Society, divididos entre oito edições lançadas semestralmente. Para efeito de delimitação do corpus de análise, se optou por mapear um total de 131 artigos, publicados nas seis primeiras edições, entre os anos de 2014 e 2016. Tal escolha se deu por meio de análise prévia do material, que constatou ser este um período de maturação das discussões sobre Big Data na revista, devido à presença inicial de abordagens epistemológicas/metodológicas, nas primeiras edições, e à publicação dos primeiros dossiês temáticos, que permitiram a identificação de múltiplas perspectivas - estudos em vigilância, meio-ambiente, arte urbana etc. - sobre a questão do Big Data nas Ciências Sociais. Atualmente, a revista tem se consolidado como referência na interface Big Data-Ciências Sociais, publicando temas especiais a partir de conferências como a *International Conference on Social Media & Society*, organizada pelo *Social Media Lab*.

### 3. O levantamento

O mapeamento, que se mostrou uma abordagem interessante para obter um panorama geral do objeto e, a partir daí, levantar diversas questões mais específicas sobre novas formas de interpretar dada literatura (PIRYANI et al., 2016; LUCHETTA, 2018), foi realizado de forma manual em seis das oito edições da publicação.

A primeira etapa do levantamento consistiu em definir os principais aspectos a serem analisados. Foram eles: autores(as), título dos artigos, palavras-chave, país de origem da instituição a que se filiam os(as) autores(as), áreas de pesquisa dos(as) autores(as), objetos de análise, principais referências utilizadas nos artigos, tipo de pesquisa (teórica, estudo de caso, etnografia, entrevista etc) publicados no período.

A partir daí o passo seguinte consistiu na investigação de cada um dos artigos das seis primeiras edições da BD&S a fim de coletar estas informações. Tais esforços envolveram *a leitura de todo o material*. No caso das referências, este levantamento contempla apenas as que aparecem mais de uma vez nos artigos, constituindo-se, assim, das referências mais importantes para a elaboração dos artigos na revista Big Data & Society. O resultado deste esforço foi compilado e representado visualmente com o auxílio do software open-source de text-mining Voyant Tools<sup>7</sup>. Os gráficos, tabelas e índices serão apresentados e discutidos brevemente nos próximos subtópicos.

---

<sup>7</sup> Disponível em: <<https://voyant-tools.org/>> Acesso em: 30/01/2019.

A partir do levantamento apresentado, alguns problemas e caminhos de pesquisa são traçados fornecendo pistas acerca das considerações sobre Big Data dentro das Ciências Sociais. Os dados a seguir constituem um panorama breve dos primeiros resultados do levantamento realizado na BD&S.

### 3.1 Geopolítico

A primeira etapa do levantamento consistiu em obter um mapeamento geográfico das publicações da Big Data & Society. Houve uma tentativa de classificar também a origem nacional de cada autor(a), todavia como nem sempre foi possível obter estes dados, preferiu-se levar em consideração apenas os países-sede das instituições de ensino dos(as) autores(as), informação esta disponibilizada pela revista. A intenção aqui era, a partir deste mapa, considerar a relação da geopolítica dos artigos com as preocupações e questões das Ciências Sociais acerca do Big Data, no contexto da Big Data & Society.



Figura 2: Origem acadêmica dos trabalhos publicados na Big Data & Society entre 2014 e 2016. As tonalidades indicam nível de presença em artigos. Tons mais escuros apontam para uma presença mais expressiva, enquanto tons mais claros indicam presença mínima. Países em tonalidade cinza clara, sem qualquer indicação, não tiveram nenhuma representação por parte de suas instituições de ensino na publicação.

O levantamento apontou que instituições estadunidenses, por meio de seus pesquisadores, encabeçam 45 dos 131 artigos publicados pela Big Data & Society no período, o que configura mais de um terço das pesquisas (34,35%). Somadas com as instituições britânicas, que representam um total de 39 artigos (29,77%), a fatia britânico-americana alcança pouco mais de 64% do material produzido na Big Data & Society. Destacam-se também a

produção por parte de acadêmicos de instituições holandesas (16 artigos no total) e canadenses (12).

Apenas quatro trabalhos partem de pesquisadores de instituições que fazem parte de países não-ocidentais: Egito, Singapura, Emirados Árabes Unidos e Japão. O levantamento, portanto, aponta para uma clara falta de representatividade de instituições de pesquisa da América Latina, África, Ásia e Leste Europeu.

### 3.2 Escolhas metodológicas

Um outro aspecto avaliado foi a escolha metodológica utilizada nos artigos da BD&S. Para esta análise, cada artigo foi analisado e classificado a partir de sua metodologia de pesquisa (ex: teórica, estudo de caso, pesquisa de campo, mapeamento de controvérsias etc.). Praticamente todos os trabalhos apresentaram claramente a metodologia no decorrer do texto. Em algumas poucas circunstâncias, foi necessário que houvesse uma categorização pessoal do tipo de método utilizado. O objetivo aqui era o de compreender as principais tendências metodológicas e quais abordagens têm sido privilegiadas - e subestimadas - na pesquisa em Big Data e Ciências Sociais.



Figura 3: Gráfico em bolha com os principais tipos de metodologia utilizada nos artigos da BD&S no período de 2014-2016. O tamanho da bolha se refere à quantidade de artigos com dada metodologia em determinado período, enquanto a progressão linear aponta para uma linha do tempo da esquerda (primeira edição) para a direita (última edição). Na imagem acima, aparecem em sequência decrescente: teórico - revisão bibliográfica (59 artigos); estudo de caso (28 artigos); comentário (11 artigos); pesquisa de campo – modalidade entrevista (9 artigos); etnografia (5 artigos) e relato de experiência (4 artigos).

Como podemos observar, o predomínio de uma metodologia teórica nos artigos da BD&S até 2016 é claro, com mais do que o dobro de artigos se valendo de estudos teóricos do que os que se basearam em estudos de caso, por exemplo. Juntas, ambas as escolhas metodológicas ultrapassam os 65% do total de publicações no período. Apesar da presença menor de outros tipos de análise no quadro geral, é possível notar no recorte temporal que métodos como pesquisa de campo, além de comentários, são mais frequentes nas edições mais recentes da BD&S, enquanto os artigos teóricos, embora ainda presentes, perdem força em comparação ao período inicial da revista (as duas primeiras edições de 2014).





Apesar do caráter pouco específico de muitas destas palavras - por exemplo, social pode se referir a inúmeras coisas - este levantamento permitiu identificar a presença considerável de temáticas como “mídias digitais”, “redes sociais”, “meio-ambiente” (11 vezes), “métodos computacionais” (10 vezes), “vigilância” (9 vezes) e “epistemologia” (8 vezes) entre os artigos. Tal achado permite inferir a formação de verdadeiros ‘campos’ de pesquisa em Big Data dentro das Ciências Sociais, como, por exemplo, os estudos em vigilância, que englobam termos como “privacidade” (6 vezes), “governança” (6 vezes), “Snowden” (5 vezes), “ativismo” (5 vezes) e “segurança” (4 vezes), entre outros<sup>8</sup>.

Além dos assuntos que mais aparecem nas palavras-chave, o mapeamento buscou apontar também a tendência de aparição dos termos no decorrer do período 2014-2016. Os termos escolhidos foram “computational” [associado à métodos]; “critical”, associado a “Critical Data Studies”; “environmental”; “epistemology”; “media”, associado a digital; “mining”, associado a data ou text; “surveillance”; e “Twitter” - todos com importante representatividade nas palavras-chave.

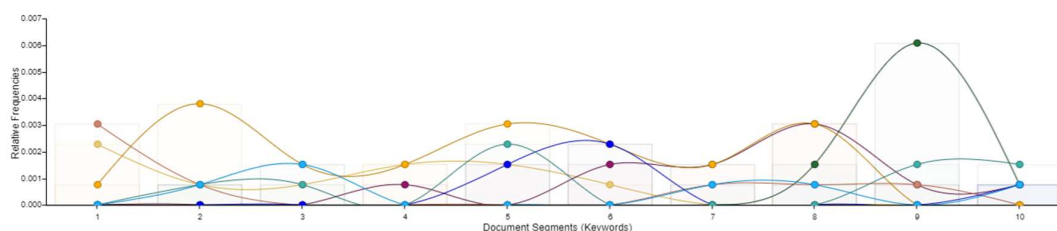


Figura 6: Para melhor compreensão: amarelo claro = computational; roxo = critical; verde = environmental; castanho = epistemology; amarelo = media; azul = mining; azul turquesa = surveillance; azul claro = twitter.

No gráfico, percebe-se que as pesquisas associadas à digital media foram mais constantes nas edições iniciais, bem como artigos relacionados à epistemologia e métodos computacionais, num período, como já apontado anteriormente, de maturação da revista, bem como dos estudos em Big Data nas Ciências Sociais. Após esta fase inicial, que comprime as duas primeiras edições, nota-se uma maior diversificação dos temas, com destaque para o crescimento de termos como vigilância e mineração [seja de dados ou textos], no decorrer das edições centrais da análise (3<sup>a</sup> e 4<sup>a</sup> edições).

Já a partir da quarta edição, e notadamente nas edições finais do escopo do levantamento (5<sup>a</sup> e 6<sup>a</sup> edições), o número de artigos com o termo meio-ambiente como carro-

<sup>8</sup> O mesmo pode ser visto também no que tange a estudos em meio-ambiente (expressões como “política de modelagem” [5 vezes] e “mudança climática” [3 vezes]), estudos em espaço e geografia (expressões como “Big Data Espacial” [4 vezes] e “rastreamento” [3 vezes]), análises de redes sociais (expressões como “mídias digitais” [20 vezes], “Twitter” [6 vezes] e “Facebook” [4 vezes]), bem como métodos computacionais (expressões como “machine learning” [8 vezes] e “mineração de textos” [6 vezes]).



Figura 7: Nuvem de palavras com o sobrenome dos(as) autores(as) mais citados(as) nos artigos da revista BD&S publicados entre 2014 e 2016.

O autor mais referenciado ao longo dos artigos da BD&S é o irlandês Rob Kitchin, professor na National University of Ireland Maynooth (Irlanda) e importante nome nos estudos em Geografia, Ciências Sociais e dados nos últimos anos. Ao todo, seu nome aparece como um importante aporte teórico em 57 vezes ao longo das publicações. Kate Crawford e danah boyd, principais pesquisadoras da Microsoft Research, aparecem 45 vezes no decorrer dos desenvolvimentos da revista, seguidas do sociólogo francês Bruno Latour (37 vezes), dos sociólogos britânicos Roger Burrows (27 vezes) e Mike Savage (25 vezes), ao lado do sociólogo norte-americano Geoffrey C. Bowker (25 vezes) e do sociólogo australiano Adrian Mackenzie (25 vezes).

É perceptível, de pronto, a relevância de teóricos associados aos Estudos em Ciência, Tecnologia e Sociedade (*Science and Technology Studies*) e à chamada Teoria Ator-Rede. Não apenas pela lista apresentada acima, mas levando em conta nomes outros com presença importante nos artigos como Michel Callon, David Lazer, Donna Haraway e Annemarie Mol. A ligação teórica entre estes autores fica clara nas informações da Figura 8.

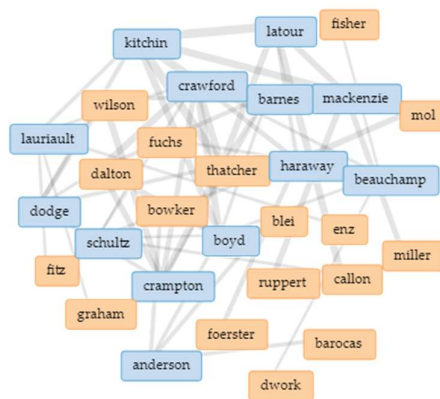


Figura 8: Ligações entre o uso de autores(as) dentro de um mesmo artigo. Linhas mais grossas apresentam autores(as) que aparecem em conjunto dentro de um mesmo artigo com maior frequência, enquanto o inverso, ou mesmo a inexistência de correlação, acontece à medida que a linha é mais fina.

Nota-se, a partir da Figura 8, que o vínculo de correlação [autores(as) que aparecem em conjunto dentro de um mesmo artigo] entre os(as) autores(as) mais citados(as), que se encontram na faixa central-superior do gráfico, é elevado, isto é, não somente aparecem com frequência nos artigos, mas em geral aparecem juntos nos textos da BD&S. Rob Kitchin, por exemplo, aparece associado com frequência elevada exatamente à(s) teóricos(as) mais

citados, como Kate Crawford, danah boyd, Jim Thatcher, Bruno Latour e Adrian Mackenzie, e vice-versa.

Tais vínculos permitem apontar, à primeira vista, a importância fundacional dos Estudos em Ciência, Tecnologia e Sociedade, em suas variadas correntes, no processo de amadurecimento da pesquisa em Big Data nas Ciências Sociais, visto que informam teoricamente uma parcela considerável dos artigos publicados entre 2014 e 2016 na BD&S.

Em linhas gerais, apesar do caráter ainda rudimentar da pesquisa, é possível afirmar o achado de importantes resultados parciais. Um deles diz respeito à presença de uma concentração euroamericana nas pesquisas investigadas. Apesar das causas de tal concentração permanecerem como objeto futuro de investigação, é notório a falta de representatividade de pesquisadores vinculados a instituições de ensino do Sul global em um debate tão importante como o Big Data na atualidade, e no contexto de uma importante publicação internacional.

A partir do levantamento prévio foi possível detectar também uma tendência maior de diversidade metodológica nas pesquisas no recorte temporal mais recente da análise. Tal achado evita generalizações sobre a pesquisa em Big Data nas Ciências Sociais, apontando para a diversidade de abordagens sobre o tema em voga atualmente.

A caracterização dos(as) principais autores(as) e de suas obras mais citadas nos artigos deve contribuir para a questão de como essas fontes fundamentam as pesquisas em Big Data nas Ciências Sociais. Afinal, a presença destas referências, seja sob a forma de rejeição ou de afirmação, certamente ajuda a moldar, de algum modo, o conteúdo do material.

### **Considerações finais**

A análise quantitativa do material publicado na BD&S entre 2014 e 2016 permitiu a identificação da origem geopolítica, das metodologias utilizadas, das palavras-chave e das principais referências usadas em cada um dos 131 artigos publicados no período. Este mapeamento é apenas um passo inicial na tentativa de apontar incômodos, reverberar potências e situar caminhos possíveis para a pesquisa em Big Data nas Ciências Sociais.

Portanto, a partir do levantamento, algumas problemáticas ficam expostas e que guiam a investigação posterior do projeto. São perguntas como:

- Considerando o domínio euroamericano na pesquisa em Big Data, em que isso influi por exemplo na escolha de referências, metodologias, objetos de análise? Como seria o Big Data e sociedade a partir do Sul global, em especial da América Latina? Que tipos de metodologias e objetos de estudo alcançariam realidades locais no que concerne ao uso e efeito dos dados na América Latina, por exemplo?

- Dentro de cada campo temático, por que algumas referências aparecem em realce em detrimento de outras? Quais as preocupações de cada um dos campos, por exemplo, os estudos em vigilância ou em geolocalização e dados espaciais?
- Que tipos de referência alimentam, por exemplo, os trabalhos teóricos? E os métodos quali-quantitativos como mapeamento de controvérsias ou etnografias?
- Avaliando todo o material levantado, e a análise crítica do conteúdo, de que forma este esforço pode colaborar com a pesquisa em Big Data nas Ciências Sociais? Quais caminhos este trabalho aponta?

Obviamente, lidar com algumas destas questões envolve um estudo mais qualitativo do que uma mera descrição quantitativa permite fazer. As próximas etapas da pesquisa ensejam por um avanço qualitativo na análise do material e nos apontamentos sobre as relações entre Big Data e Ciências Sociais.

### Referências

AKOKA, Jacky; COMYN-WATTIAU, Isabelle; LAOUFI, Nabil. Research on Big Data - A systematic mapping study. *Computer Standards & Interfaces*, v.54, n.2, p. 105-115, 2017.

BOYD, danah; CRAWFORD, Kate. Critical questions for Big Data: provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, v.15, n.5, p. 662-679, 2012.

CHEN, Shu-Heng (Ed.). *Big Data in Computational Social Science and Humanities*. Cham, CH: Springer, 2018.

\_\_\_\_\_; YU, Tina. *Big Data in Computational Social Science and Humanities: an introduction*. In: CHEN, Shu-Heng (Ed.). *Big Data in Computational Social Science and Humanities*. Cham, CH: Springer, 2018.

CONFORTO, Edivandro Carlos; AMARAL, Daniel Capaldo; SILVA, Sérgio Luis da. Roteiro para Revisão Bibliográfica Sistemática: aplicação no desenvolvimento de produtos e gerenciamento de projetos. In: Congresso Brasileiro de Gestão de Desenvolvimento de Produto - CBGDP, 8., 2011, Porto Alegre-RS. *Anais...* IGDP, 2011.

DALTON, Craig; THATCHER, Jim. *What does a Critical Data Studies look like and why do we care?* - Society and Space, 2014. Disponível em: <<http://societyandspace.org/2014/05/12/what-does-a-critical-data-studies-look-like-and-why-do-we-care-craig-dalton-and-jim-thatcher/>> Acesso em: 16/12/2018.

FAZEKAS, Mihály. *The Use of 'Big Data' for Social Sciences Research: an application to corruption research* - SAGE Research Methods Cases, 2014. Disponível em: <<http://methods.sagepub.com/case/big-data-for-social-sciences-research-an-application-to-corruption-research>> Acesso em: 16/12/2018.

FOSTER, Ian; GHANI, Rayid; JARMIN, Ron S.; KREUTER, Frauke; LANE, Julia (Eds.). *Big Data and Social Science: a practical guide to methods and tools*. Boca Raton, FL: CRC Press, 2017.

KITCHIN, Rob; MCARDLE, Gavin. What makes Big Data, Big Data? Exploring the ontological characteristics of 26 datasets. *Big Data & Society*, v.3, n.1, p. 1-10, 2016.

LANEY, Doug. *3D data management: Controlling data volume, velocity and variety*, 2001. Disponível em: <<http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf>> Acesso em: 18/11/2018.

LUCHETTA, Sara. Literary Mapping: At the Intersection of Complexity and Reduction. *Literary Geographies*, v.4, n.1, p. 6-9, 2018.

MILAN, Stefania; TRERÉ, Emiliano. *Big Data from the South: The Beginning of a Conversation We Must Have*. 2017. Disponível em: <<https://ssrn.com/abstract=3056958>>. Acesso em: 16/12/2018.

PIRYANI, Rajesh; MADHAVI, Devaraj; SINGH, Vivek Kumar. Analytical mapping of opinion mining and sentiment analysis research during 2000–2015. *Information, Processing and Management*, v.53, n.1, p. 122-150, 2017.

SILVA, Guilherme Cavalcante. Para Repensar Comunicação e Big Data: situando caminhos a partir da edição especial do Journal of Communication. *Revista do Edicc*, v.5, p. 107-118, 2018.