

18º Congresso Brasileiro de Sociologia

GT 39 – Sociologia Digital

**Métodos quantitativos e pesquisas de levantamento nas
Ciências Sociais e na Sociologia Digital brasileira**

Elismênnia Aparecida Oliveira (UFG)

Matheus Guimarães Mello (UFG)

26 a 29 de julho de 2017, Brasília (DF)

Resumo

Partindo da perspectiva teórico-metodológica documental, neste artigo, apontamos problemáticas metateóricas acerca da sociologia digital demonstrando que as pesquisas de levantamento quantitativo são antecessoras a ela, ao mesmo tempo em que a sociologia digital é um novo campo social de pesquisa que requer novas problemáticas, técnicas e metodologias. Assim, a partir da historicização da sociologia digital, em um primeiro momento, tecemos considerações sobre a construção de categorias e variáveis, para, em um segundo momento, discutir o acesso e a liberdade de produção de pesquisas com a utilização de *big data* no Brasil. Utilizando as vertentes epistêmicas da sociologia do conhecimento, história e filosofia da ciência, adentramos temas referentes a relações de poder-saber sobre: acesso, manutenção e formação dos bancos de dados; atrelamento das pesquisas de levantamentos, e digitais, a uma ciência que produz verdade; liberdade de acesso a *big data* contraposta pelo tratamento dos dados mediante utilização de programas e linguagens especializadas, presentes em larga escala apenas nos grandes centros de produção tecnológica, econômica e política e as; dificuldades de formação de cientistas sociais nas áreas digitais e quali-quantitativas. Por fim, apresentamos a necessidade de uma transdisciplinarização entre ciências sociais, ciências da comunicação, computação e sociais aplicadas para o desenvolvimento pleno da sociologia digital no país.

Introdução

Seja para construir mapeamento de identidades regionais ou criar previsões de eleições, medir o consumo e uso de produtos e bens, gerenciar gastos públicos em distintas regiões ou elaborar parecer sobre relações sociais de violência, escolaridade, saúde, lazer, sexualidade, dentre vários e distintos temas, os recenseamentos e levantamentos populacionais têm sido encarados como potenciais para a elaboração de diagnósticos, inferências e interferências sociais. Mas, se por um lado, o uso de estatísticas por levantamento de dados, resumidas e apresentadas em gráficos e tabelas, seja recorrente para fundamentar a argumentação pública, midiática, empresarial e política de opiniões e decisões, por outro, ela não chega a ser, nas ciências sociais no Brasil, uma das técnicas mais apreciadas e utilizadas em larga escala.

Junto a problemáticas de execução, preparação e ida a campo para coletar dados nas pesquisas quantitativas, temos um consolidado interesse nas

metodologias qualitativas, ao longo da história das ciências humanas e pesquisa social, em que chama atenção a baixa produção teórica nacional sobre metodologias quantitativas. As buscas nos portais de produções acadêmicas Redalyc e Periódicos Capes¹ tem poucos resultados de artigos e pesquisas, e apontam que as principais publicações nas áreas estão na demografia, saúde, ciências ambientais e por último, ciências sociais aplicadas. A quantidade de textos na sociologia, ainda que sejam metateóricos, sobre o tema, não chega a três artigos, confirmando a pesquisa de Marília Ramos (2013), com levantamento de 1990 até 2000, que apresentava resultados tão baixos quanto os atuais. Em relação a produção sobre *big data* e sociologia digital o cenário nacional é o mesmo em relação a ausência de textos. A maior parte das produções sobre sociologia digital estão localizadas na região do Vale do Silício, escritas em inglês, e os poucos textos em português não somam mais que 10 produções.

No entanto, essa baixa quantidade de publicações e pesquisas quantitativas faz parte de indicadores multifatoriais que não demonstram apenas a ‘ausência’, ou falta de interesse, já que ao contrário, segundo Daniel Macfarland, Lewis Kevin, Amir Goldenberg (2015) e Leonardo Nascimento (2016) estamos entrando na era da sociologia digital, desde a inserção de *big data*² nas ciências humanas. Assim, como apontam algumas vertentes epistemológicas, ao contrário da falta de status das metodologias quantitativas, existem problemas estruturais, tais como: a) dificuldades de acesso a informações, programas e técnicas; b) manipulação de indicadores sociais; c) hegemonia dos métodos quantitativos, sustentada por sua vinculação a

¹ O Redalyc <www.redalyc.org/> (Rede de Revistas Científicas da América Latina e do Caribe, Espanha e Portugal) e o Periódico Capes <www.periodicos.capes.gov.br/> são importantes sites agrupadores de periódicos, livros e revistas de seus países e redes latino-americanas, no caso da Redalyc, e internacionais no caso do Periódico Capes. A pesquisa realizada em ambos aconteceu em agosto de 2016 com meu login da universidade por busca das palavras-chave: métodos quantitativos; estudos quantitativos; estudos de levantamento; e demografia vinculados a ‘sociologia e Brasil’. As buscas resultaram em duas publicações na sociologia, ambas de autoras e autores brasileiros, as demais são da saúde, demografia, ciências ambientais e ciências sociais aplicadas.

² O termo *big data* vem das ciências da informação e designa uma “alta quantidade de dados” presentes em servidores de redes de computadores, que utilizam um sistema operacional e estão interligadas em rede, e que muitas vezes, têm quantidades de informação que estão para além dos programas e aplicativos de processamento utilizados historicamente nas metodologias quantitativas, incentivando o uso de novas linguagens de programação e computação para a execução de pesquisas.

pesquisas empíricas, ou do que se considera 'mundo real' e verificável, que afetam desde o desenvolvimento de metodologias quantitativas até a solidificação da sociologia digital.

Pensando essas questões, nossa proposta neste artigo é argumentar sobre a construção e uso das pesquisas de levantamento, e sua relação posterior com a sociologia digital, em relação ao acesso, construção de categorias e variáveis a partir de relações sociais de poder presentes na produção de conhecimento para no segundo momento discorrer sobre possibilidades que têm sido levantadas no campo metateórico dessas áreas. Nesse sentido, daremos ênfase ao contexto e acesso aos dados junto aos métodos e técnicas de pesquisas de levantamento em suas definições e construções científicas, utilizando para tanto a história e filosofia da ciência.

Metodologias quantitativas nas ciências sociais e sociologia digital no Brasil

Yvonna Lincoln e Egon Guba (2006) constroem uma sequência sobre paradigmas desenvolvidos ao longo dos séculos da história da ciência, em que apresentam quadros comparativos entre metodologias, técnicas e propostas ontológicas, morais e éticas, dentre outras categorias, que, embora surjam em momentos específicos, não são apresentados como lineares, mas confluentes, na qual podemos inferir que a escolha pelas vertentes teórico-metodológicas acontece graças ao pertencimento ou a iniciação de estudantes em uma comunidade acadêmica, ou a uma universidade, a linhas de pesquisa, temas, linhagem de pesquisa, orientação, formação e até por país de pertencimento, e que mesmo as mais criticadas como o 'positivismo' não estariam 'completamente ultrapassadas', mas teriam alguns de seus aspectos ainda em uso em distintas vertentes e escolas.

Nesse contexto é possível entender a coexistência de várias vertentes e técnicas distintas, ao mesmo tempo em que é possível mapear as relações de 'maior uso', 'poder e prestígio' de algumas metodologias em relação a outras. No Brasil, como aponta Marília Ramos (2013), a partir de sua pesquisa sobre teses, no fim da década de 1990, existia uma preferência pelas metodologias qualitativas nas ciências sociais quando comparadas a pesquisas quantitativas, e ainda nos anos 2000 a oferta de formação de cursos voltados aos métodos quantitativos ainda é a

USP e a UFMG, sendo o segundo mais reconhecido, dado como curso de verão, fundamentado no ensino de metodologias quantitativas dos EUA, e mantido sem financiamento regular da Capes e do CNPq (RAMOS, 2013). A relação com estudo de *big data* também não se difere, já que existem poucos laboratórios e núcleos de pesquisa com atividades sendo desenvolvidas, nessas mesmas universidades que desenvolvem metodologias e ensino quantitativos, USP e UFMG. A diferença são pesquisadores atuantes, com produção de material e pesquisa, na região centro-oeste e nordeste, mas ainda sem disponibilidade de ‘cursos exteriores’ sobre o tema. Assim, a sociologia digital segue sendo ‘urgente’, mas pouco escrita e registrada.

Considerando alguns manuais sobre o tema, Uwe Flick (2013), Floyd J. Fowler Jr (2011), podemos apontar que, apesar da tendência da maior parte das propostas de pesquisa serem da utilização do método de triangulação, ou seja, para o uso de metodologias *quali-quantitativa*, existem algumas complicações para a execução das pesquisas quantitativas, que estão reproduzidas no estudo de sociologia digital, tal como mapeia Marília Ramos (2013), Louis Rea e Richard Parker (2000) como: a) o uso de softwares para tabulação e criação de dados; b) o estudo e a constituição de uma amostra válida e significativa; c) a aplicação de entrevistas nos grupos de amostragem e o estudo de probabilidade causal; d) a escolha de algumas causas e a combinação delas em detrimento de outras, já que a multicausalidade não pode ser acessada (REA; PARKER, 2000; RAMOS, 2013). Esses fatores fazem com que a produção de pesquisas quantitativas possa ser, portanto, a princípio, dispendiosa, requerer formação específica e profissional no uso de softwares e formação de equipes para execução, o que acaba restringindo o uso para apenas algumas técnicas quantitativas. Todo esse arsenal de problemáticas das pesquisas quantitativas persegue a sociologia digital, que não deixa de ser, sob vários aspectos uma sociologia de metodologia quantitativa.

Portanto, se esses problemas acontecem desde os primórdios da utilização de metodologias quantitativas, é ainda mais complexo no estudo com *big data*, em que a quantidade de informações produzidas e acessadas sequer consegue ter instrumentos de armazenamento e análise que ‘estão a disposição’, mas precisam ser ‘literalmente construídos e inventados’ a partir do desenvolvimento de

programas e algoritmos. E nos dois casos, tanto nas metodologias quantitativas quanto para sociologia digital, a maior parte das dificuldades ainda passam pelo crivo social, não são apenas estruturais, ao contrário, a tradição de escolha por metodologias quantitativas e qualitativas têm peso maior na escolha de formação das pesquisadoras e pesquisadores, e no Brasil, Marília Ramos (2013) aponta que a tradição brasileira concentra na ciência política um maior uso de técnicas e análises quantitativas baseadas em estatísticas e levantamentos, e tem seguido essa tradição na utilização de estudos sobre *big data*.

Na sociologia, os estudos de levantamento aparecem de forma interdisciplinar, com maior visibilidade na demografia e estatística, como acontece nas análises produzidas pelas instituições estatais IBGE e IPEA³. Assim, ainda que no Brasil exista uma larga produção de bancos de dados, a partir da década de 1990, com ênfase nos censos produzidos pelo IBGE, utilizando a aplicação de questionários em domicílio, o uso desses dados em pesquisas não tem uma produção de análise amplamente analisada para além do próprio IBGE e do IPEA, ambos órgãos institucionais que realizam os censos e análises por meio de equipes multidisciplinares, estruturadas a partir da geografia, demografia, e estatística e conjuntas as ciências sociais e ciências sociais aplicadas.

O fato de as pesquisas quantitativas no Brasil terem seu principal polo de produção em órgãos governamentais, traz um indicativo importante sobre os usos dessas metodologias e sua localização na demografia e economia, como formas não só de análise e desenvolvimento de políticas públicas, mas também de controle social desde o período colonial, na história do IBGE (IBGE, 2016), e a partir da ditadura militar, com o histórico de criação do IPEA (IPEA, 2016). No entanto, ainda que não seja de interesse a construção de problemáticas e hipóteses das ciências sociais, o controle populacional, ou o controle de amostras na representação de

³ No Brasil os organismos estatais de medição da população existem desde a Diretoria Geral de Estatísticas, criada em 1871, que foi extinta em 1934, para a criação em 1936 do INE, Instituto Nacional de Estatística, tendo incorporado o Conselho Nacional de Geografia se tornando o Instituto Brasileiro de Demografia e Estatística <<http://www.ibge.gov.br/>> (IBGE) no mesmo ano. O Instituto de Pesquisa Econômica Aplicada <<http://www.ipea.gov.br/>> (IPEA), por sua vez, vem substituir em 1964 o EPEA, Escritório de Pesquisa Econômica Aplicada criado em 1960. O IPEA é vinculado ao Ministério do Planejamento, Desenvolvimento e Gestão, e tem por descrição institucional pesquisar e analisar políticas públicas e ações governamentais por meio da análise e realização de pesquisas e estudos.

comensurabilidade de perguntas e hipóteses (RAMOS, 2013; REA, PARKER, 2000), são essas características, a) possibilidade de controle; b) verificação e a; c) comensurabilidade, que garantem o prestígio das metodologias quantitativas para além das ciências sociais e que preparam o cenário acadêmico para a entrada da sociologia digital.

A sociologia digital aparece como relevante nas ciências sociais como vinda do grupo de sociologia quantitativa, uma sociologia mais próxima a realidade social, com a valorização e popularização da alta quantidade de informações por demonstração de domínio e análise de dados, mas faz surgir novos objetos e demandas de estudos a partir do 'sistema de informações de redes sociais que conectam pessoas e grupos'. A relação com 'quem tem o controle dos dados' está não mais, apenas, nas 'mãos do estado', mas como aponta Sophie Mützel (2015), nas 'mãos de empresas privadas' que têm se apropriado do uso da internet, e da *web*, na construção de redes e conexões entre pessoas, para a venda de informações. Será então, mediante esses problemas, que na sociologia digital a crença na possibilidade de 'medir e acessar a realidade' está realmente assegurada? A sociologia digital espelha os problemas da sociologia quantitativa?

Entre a comensurabilidade e a criação de verdades: métodos quantitativos na sociologia digital

Tendo como base o princípio de mensuração dos fenômenos sociais, de acordo com Marília Ramos (2013) e Fred Kerlinger (1980) a ciência empírica quantitativa se preocupa com resultados gerais e coletivos, sendo contextual e multinível. Desde o positivismo, com Auguste Comte, que o empirismo vinculado ao comensurável, segundo Kerlinger (1980) e Lincoln e Guba (2008), pode ser aplicado às ciências humanas garantindo a verificação e falseamento de hipóteses por meio do estudo de probabilidades estabelecidas nas amostras a partir de relações de causalidade entre as variáveis, tal como era usada nas áreas das ciências exatas e biológicas, para medir e analisar percepções, valores sociais, e interesses.

Longe de apostar na previsão de valores e éticas, essas autoras e autores pontuam que a partir da formulação criteriosa do problema de pesquisa, e das hipóteses, é possível construir métodos de análise nas amostras, como

questionários e grupos de discussão, que consigam, pautados em um sólido referencial teórico, inferir sobre qualquer tema social de forma exitosa, ou seja, apontando generalidades a partir de causalidades. Assim, as pesquisas empíricas saem dos horizontes das áreas exatas, e passam a ser, nas ciências sociais, sobre vivências e percepções das pessoas, bem como sobre relações sociais entre países, organizações, famílias e agrupamentos sociais (RAMOS, 2013).

A diferença no surgimento da sociologia digital é o estudo em amostra de “comentários, curtidas, reações e compartilhamento de links”, por exemplo, em que a forma de ‘acessar os discursos e interações dos sujeitos para inferir sobre eles’ não necessita da interação direta com as pessoas pesquisadas. O acesso a ‘rede pública’, que as pessoas compartilham, já torna possível realizar várias investigações com distintos objetivos. O trabalho com *big data* em que se estude redes como Facebook ou Twitter tem na ‘mensuração de comentários, curtidas e reações por emoticons a possibilidade de “mensurar sentidos e reações” diante de variados temas, mas também pode-se estudar a partir de ‘curtidas, compartilhamentos de links e respostas a jogos de personalidades’ relações sobre ‘raça-cor-etnia e sexualidade’ que serão usadas para ‘eleger propagandas, produtos, e posicionamentos políticos no histórico de acesso’ da web, tal como foi executado na pesquisa realizada por Eytan Bakshy, Solomon Messing e Lada Adamic (2015) que estudando usuárias e usuários do facebook a partir de suas curtidas, interações e comentários, propuseram uma análise social sobre posicionamentos políticos, consumo e interação política.

Nesse estudo de caso do Facebook é possível perceber que algumas formas de executar pesquisa na sociologia digital se aproximam dos métodos quantitativos, seja pela ‘descrição identitária dos grupos’, seja por ‘como essa descrição é feita a partir da mensuração de dados. Essa dinâmica de ‘mensurar’ não é tão distinta do primeiro recenseamento no Brasil, em 1872, em influência da participação de brasileiros no Congresso Internacional de Estatística, realizado nesse mesmo ano (PETRUCCELLI, 2013), que tinha como uma das principais funcionalidades ‘descrever a população brasileira’ em relação aos marcadores de raça/cor/etnia. Os estudos identitários por sua vez, feitos em relação aos big datas, seriam feitos utilizando fotos, compartilhamento de postagens e várias outras

informações que cumprem o papel da ‘autoclassificação’, mas ainda dependem do aprofundamento teórico-metodológico das pesquisadoras e pesquisadores, bem como de seus interesses e envolvimento com a pesquisa desenvolvida.

E se ao longo dos séculos os marcadores identitários, utilizados nos recenseamentos, foram modificados para ‘uma maior precisão dos dados’, as pesquisas quantitativas foram construídas, desde que assumem uma relação com a investigação empírica, com uma identificação maior com valores científicos largamente consolidados tais como: “verificação; neutralidade; objetividade e representação da realidade” que as metodologias qualitativas. Essas mesmas características, por outro lado, nem sempre são relacionados às pesquisas qualitativas que são vistas, muitas vezes, como aponta Flick (2013) e Ramos (2013), aprofundadas sobre um circuito reduzido, enquanto as quantitativas conseguiriam abranger um raio maior de análise social em relação a quantidade e influência coletiva. Assim, outro aspecto que a sociologia digital tem herdado da sociologia quantitativa é o processo de ‘trazer verdade a partir dos dados’, como se existisse uma correlação simétrica entre “quanto maior a quantidade de dados, maior a possibilidade de produzir verdade sobre um tema”.

Sendo as ciências sociais baseadas na construção de teses sobre argumentos, o uso de estatísticas e resultados de levantamentos são prestigiados e fundamentadores de ‘argumentação’ considerada relevante. Não é por acaso que a mídia, as empresas, e os governos os utilizam com frequência. No livro sobre o tratado da argumentação de Chaim Perelman e Lucie Olbrecht-Tyteca (2014), em que estudam a fundamentação dos argumentos para sua validação e convencimento social de confiabilidade, vários pontos sobre ‘quantidade’ são pontuados, tais como o que chamam de ‘lugares de quantidade’, em que a argumentação se estabelece pela premissa forte da relação entre recorrência-normal-normalidade; ‘argumentos baseados na estrutura do real’ em que ‘recorrência, generalidades e causalidades’ são vinculadas a fundamentação da argumentação; e argumentos fundamentados em ‘probabilidades’, que são construídos com base no uso de estatísticas. Em uma produção semelhante sobre os usos do argumento, Stephen Toulmin (2006) pontua que a argumentação é baseada em probabilidade e tem layouts baseados em dados e generalizações que,

embora tenham perigos e falhas, são socialmente reconhecidos e utilizados.

Segundo Fowler (2011), Stephen Toulmin (2006) e Chaim Perelman, Lucie Olbrechts-Tyteca (2014) é possível pensar a relação entre teoria e metodologia quantitativa a partir do prestígio e das falhas das técnicas quantitativas não apenas pela importância da teoria para análise social dos dados, mas também para a composição dos métodos. Nesse sentido, entra em questão o convencimento e proposição de verdade que estatísticas trazem desde seu histórico de confiabilidade construído a partir das ciências exatas. A própria concepção de 'dado' para referir aos resultados da pesquisa indicam a crença no 'real acessado e quantificado', o que demonstra que o grande público fora das ciências sociais, e por muito tempo dentre cientistas sociais, que têm acesso aos gráficos e aos resultados dos levantamentos, não imagina que os mesmos tenham margem de erro, questionários com perguntas capciosas e truncadas, ou que as amostras não tenham representação real e total da população sobre o tema e sim uma 'grupo de amostra', uma quantidade específica acessível, criada e escolhida.

Nesse contexto, embora Kerlinger (1980) e Lincoln e Guba (2006) apontem que os métodos quantitativos têm um largo processo epistêmico de crítica, por vários outros paradigmas epistemológicos das teorias críticas, construtivismo e participativismo (GUBA, LINCOLN, 2006) e autocrítica científica sobre si, ao repensar os critérios de 'neutralidade, objetividade, e falseamento' (KERLINGER, 1980), a construção histórica do empirismo e das metodologias quantitativas ainda precisam ser estudadas de forma mais ampla, tanto pelas utilizações para manipulação quanto pela 'falta de acesso a dados'. Na sociologia digital essa discussão permanece tanto em relação a manipulação e falta de acesso na íntegra aos dados, já que é impossível ter acesso a bancos de dados de empresas como Google, Facebook, e Twitter da mesma forma que pesquisadoras e pesquisadores que trabalham nessas empresas têm. Podemos perceber que a sociologia digital abrange várias possibilidades em relação a 'produção de verdade, construção de amostra, confiabilidade em bancos de dados', herdada das 'sociologias quantitativas', só que trazendo dificuldades mais escancaradas e contraditoriamente abrangendo o campo de pesquisa social. É pensando nessas contradições e nessa 'herança' que discutimos algumas das principais limitações do estudo de big datas,

e da sociologia digital brasileira.

O acesso aos dados digitais e a sociologia digital no Brasil

As problemáticas dos métodos quantitativos apresentadas até aqui possuem relação com a sociologia digital por a anteceder nos aspectos de 'técnicas desenvolvidas ao lidar com alta escala de dados empíricos e coletar, construir, esses dados'. Mas ao mesmo tempo, a sociologia digital traz demandas específicas que vem complexificar ainda mais as questões de pesquisa com *big data*. Embora muitas categorias da análise da história da ciência, tal como os conceitos de paradigma, comunidades científicas, técnicas de pesquisa e epistemologia, sejam fundamentalmente importantes, na medida em que novas formas de armazenamento social da produção de conhecimento surgem, são exigidas e criadas, novas formas de análise dessa produção, com novas categorias conceituais que por sua vez representam as mudanças sociais em curso.

Tal como apontam Dominique Vinck (2013), Sophie Mützel (2015), Serge Noiret (2015) Daniel Macfarland, Lewis Kevin, Amir Goldenberg (2015) e Leonardo Nascimento (2016) a produção de conhecimento em ciências sociais, principalmente na sociologia, tem passado por uma estrondosa mudança na produção de pesquisas graças a novas formas de armazenamento de conhecimentos e informações pessoais em *big data*. Agora não só 'realizamos os recenseamentos' e lidamos com os problemas em relação às categorias e pesquisa de coleta em campo, mas lidamos com novas linguagens, formas de produção e armazenamento, nunca antes conhecidas. Desde bibliotecas digitais até 'comunidades virtuais de interesses', estamos diante de um alto armazenamento de 'formas de comunicar e ser' (NASCIMENTO, 2016).

A sociologia digital surgiu nesse contexto para visibilizar campos de pesquisa próprios, não só completamente novos (NASCIMENTO, 2016, MÜTZEL 2015), mas também renovados. Estamos em um momento em que alguns objetos de pesquisa passam a ser alocados na produção digital sendo ampliados e rearticulados graças a eficácia do arsenal da análise sociológica às outras ciências (MÜTZEL 2016). Um exemplo dessa rearticulação pode ser pensado com o estudo

da 'produção de conhecimento', seja ela científica, tendo em vista artigos, armazenamento de obras de arte, textos e livros, ou cotidiana, na forma de posicionamentos políticos em publicações, memes, e compartilhamento de notícias e links', ou seja, com a discussão e o desenvolvimento de técnicas de estudo sobre o digital, a produção de conhecimento pode ser acessada e literalmente configurada enquanto uma 'rede social de ideias e pessoas, autorias, em conexão'.

Além disso, os vários algoritmos mediando a vida cotidiana permitem entender que *softwares* e programas podem ser 'enviesados, injustos e discriminatórios' (MÜTZEL, 2015) demonstrando as relações de poder entre 'quem detém e usa esse conhecimento para gerir a vida pública' e entre 'quem utiliza esses mecanismos' no cotidiano. No caso do Brasil, essa rede de informações, objeto da sociologia digital, pode ser reconhecida pelo próprio sistema de cadastro de docentes e estudantes em sistemas de 'plataformas de revistas, *blogs*, vídeos, e livros' que por sua vez formam sistemas de publicações, citações, atividades realizadas, ou seja, é possível verificar e inferir sobre o que Thomas Kuhn (2001) nomeou de comunidades acadêmicas e até sobre a produção de paradigmas ao longo do tempo por meio do 'mapeamento e gerenciamento' de informações digitalizadas com *hiperlinks*. Algo impensável por métodos quantitativos até então conhecidos.

Junto ao armazenamento da 'produção de conhecimento' vários outros estudos podem ser desenvolvidos a partir do acesso a 'bancos de dados via *web*' vinculados a áreas de saúde, trabalho e lazer, e tendo em vista interesses e buscas para consumo. A partir do arsenal da sociologia digital com o desenvolvimento de métodos digitais para extrair e armazenar dados, para rastrear e construir relações entre buscadores e *downloads*, temos uma ampliação tanto na metodologia quantitativa quanto na produção da sociologia do conhecimento. Pelo surgimento da sociologia digital é possível 'identificar' que a produção de conhecimento, dentre ela a acadêmica, pode ser pensada como constituinte de uma rede numa postura que reorganiza o 'estudo da produção de conhecimento' e renova a análise sobre o pensamento social. Tal como aponta Sophie Mützel (2016), a análise com *big data* de forma qualitativa ou quantitativa, traz a sociologia em especial, uma 'utilidade social' que antes era invisibilizada, a de inferir sobre dados, posicionamentos e

ideias, ao mesmo tempo em que ‘aprofunda olhares sociológicos’ sobre as interações sociais e relações de poder-saber.

Nesse contexto, se por um lado, as ciências sociais se legitimaram enquanto ciência desde o reconhecimento da produção demográfica, estatística e da metodologia quantitativa em geral, por outro, o surgimento do *big data*, sobre vários e distintos temas, bem como o acesso a ‘informações largamente compartilhadas em armazenamento de nuvens’, a questão metodológica e epistemológica da sociologia quantitativa não passa mais apenas pelas problemáticas próprias das ciências sociais e aplicadas, mas apresenta um momento único que entrelaça a produção científica das ciências sociais às mídias digitais, engenharias e a computação (NOIRET, 2015; ZIEWITZ, 2015; MACFARLAND, KEVIN, GOLDENBERG, 2015) em um complexo processo de controle/gerenciamento e armazenamento de pessoas, informações e conhecimentos.

É preciso considerar para esse momento que todas as pessoas estão cotidianamente sendo submetidas a algoritmos (ZIEWITZ, 2015) a instrumentos programados para gerir aspectos de suas vidas, seja no trabalho, no acesso a internet, no trânsito, na escola, enfim, em todos os ambientes de convivência existem ‘programas e linguagens’ sendo utilizadas para, ‘a partir da coleta de dados’ fazer inferências e às permitir ‘caminhar, consumir, entrar e sair’. Neste momento, portanto, aconteceu, e fica cada vez mais evidente, a ‘modificação da cultura’ e do acesso ao cultural e as representações culturais (ORTIZ, 2002; NORIET, 2002), já que temos acesso a ‘essa cultura’ de ‘qualquer lugar’ e várias vezes de ‘qualquer forma’, mas é também o momento em que estamos, majoritariamente, sem acesso público aos meios de produção dessas linguagens. Não temos, entre nós, o mesmo tipo de acesso a internet, nem no país, nem mundialmente. Os mapas de cabos de conexão submarinos de *web* que ‘ligam o mundo’, atravessam em maior escala dos EUA e da Europa para países de África e América Latina e Insular, ou seja, ‘dos países colonizadores para os países colonizados’ (PIMENTA, 2016), como se o desenvolvimento da internet em si, bem como o desenvolvimento das *big data*, estivesse atrelada à relações políticas do século XVIII.

Os problemas de “acesso aos dados que ficam apenas com órgãos estatais ou com as empresas privadas que detêm as informações compartilhadas em redes” agora é também sobre ‘saber para acessar’. Estamos, ainda que o estudo de tecnologias digitais tenha aumentado a ponto de trazer para a linguagem cotidiana as discussões sobre algoritmos (DOURISH, 2016; ZIEWITZ, 2015), diante de uma velha questão: as relações de poder presentes na produção de conhecimento, que geram as formas de poder-saber (QUIJANO, 2005). Mesmo com a possibilidade de acesso, pela vivência digital, continuamos a produzir linguagens e dados sobre o cotidiano, desta vez o computacional e tecnológico, que estão completamente distantes da maior parte da população, mas presente na maior parte de atividades que executam.

Essa questão, acompanhada e justificada na criação de ‘especialistas tecnocratas’, reproduz o que Aníbal Quijano (2005) nomeou de ‘relações de poder-saber’ baseadas numa construção da ‘ciência como verdade’ utilizada em processos de governabilidade e controle social. Se durante as décadas de 1990 e 2000 a preocupação epistêmica dos estudos culturais e descoloniais eram com o processo de imposição da cidadania apenas para os ‘letrados’ por meio do ‘letramento da cidade’, os problemas de tradução, a burocratização centrada na escrita que produziu leis, normas, documentos e ‘conhecimento verdadeiros’ em livros, áudios e outras mídias (STEPHAN, 1996; CASTRO-GOMEZ, 2005; SPIVAK, 2002) agora, o problema não é necessariamente o acesso à informação, mas o acesso ‘às formas em que a informação tem sido produzida e armazenada’, bem como aos ‘algoritmos’ que, como apontam Malt Ziewitz (2015) e Paul Dourish (2016), estão presentes não só nos celulares e computadores, mas no comércio, transporte, na vida pública política de acesso à saúde, ao conhecimento e a educação, e a maior parte da população segue sem saber sequer acessar esses dados, ou que estão participando da coleta dos mesmos, agora a questão de letramento passa pelo cunho digital.

O que a sociologia digital tem a dizer sobre acesso à produção de conhecimento?

Neste artigo, procuramos trazer algumas considerações críticas e

desdobramentos dos estudos de levantamento nas ciências sociais no Brasil. A princípio mapeamos as relações de prestígio da pesquisa empírica para num segundo momento apontar algumas críticas vindas de dentro das vertentes metateóricas das ciências sociais e sua correlação com os estudos da sociologia digital, um campo novo de pesquisa. Apontando a relação de ‘verdade e estudo do real’ pontuamos que os estudos de levantamento tiveram parte de seus ideais criticados a ponto de fazer com que fossem repensados a partir de contextos sociais e variáveis, mas que seguem sendo ‘de acesso e manuseio exclusivo’ para alguns grupos e saberes embora sejam amplamente utilizados.

Ressaltamos ainda que um estudo mais detalhado com vertentes descoloniais e feministas sobre o tema ajudariam a questionar posturas solidificadas sobre as metodologia quantitativas e sobre a herança de algumas dessas posturas na sociologia digital. Apresentamos por fim alguns desafios das ciências sociais, da sociologia digital, e das pesquisas de métodos digitais, começando pela inserção da formação, desde a graduação, ampla e transdisciplinar com metodologias quantitativas e digitais de áreas como a informática, o sistema de informação, e a comunicação nas ciências sociais (NASCIMENTO, 2016). Atrelado ao desafio da formação temos a ausência de acesso aos dados, que por sua vez, estão correlacionados a uma ‘produção constante de patentes de conhecimento’, ou de ‘patentes de detentores de *big datas*’. Como apresenta Dominique Vinck (2013), é preciso pensar no problema de detenção de ‘meios de saber’ acoplados com estratégias de controle social por ‘venda desse conhecimento’. O completo controle da *google* a várias informações, sites, imagens, documentos históricos, dentre outros, não é ‘neutro e objetivo’ a vida das pessoas.

O preço para o acesso a *softwares*, as vendas casada entre sistemas operacionais em computadores e a construção de uma rede de ‘mídias’ obrigatória, que sem permissão, armazenam dados pessoais, e incentiva compartilhamento de dados, tem sido a face mais humana e cotidiana da interação com o mundo digital das pesquisas sociais. Somado a isso temos nichos restritos de ‘tecnologias livres’ baseadas na criação de redes de criação e compartilhamento de conhecimento tecnologias e mídias digitais que não chegam a ‘universidades, e a grupos de pesquisa’ e são ainda mais restritas. Ou seja, a falta de formação dos estudantes de

ciências sociais reflete a forma de informação da maioria da população do globo sobre a *web* e seus componentes de controle social vinculados aos governos, ao comércio e a cultura.

Não adianta ter uma ‘história digital’, seguramente armazenada, se ela não ‘será acessada por todas as pessoas’, nem ‘conhecida’ em como foi organizada, nomeada e escolhida. Não é ao acaso que a produção de uma ‘sociologia digital, de fato transdisciplinar e com pesquisadores’ com alta produção e desenvolvimento de área temática estejam nos Estados Unidos da América, no Vale do Silício, um polo de produção empresarial e computacional, isso condiz com os processos de colonização que ainda atuam na distribuição e produção do poder (VINCK, 2013; PIMENTA, 2016). No Brasil, embora ainda não contestado, o ‘locus de produção’ de metodologias quantitativas, como levanta Marília Ramos (2013), com ensino de métodos que poderiam ser usados na sociologia digital estão localizados na região sudeste, Minas Gerais e São Paulo, que também reproduzem a dinâmica de ‘pólos comerciais e de produção de conhecimento’.

O resultado desses dois grandes desafios, formação e patenteamento de dados, de algoritmos e de sistemas de informação, tem reverberado na baixa quantidade de produções da sociologia digital e nas dificuldades de seu desenvolvimento no país. Existem raras publicações de pesquisas que compartilham ‘algoritmos’ produzidos para pesquisar determinadas áreas sociais, tal como acontece com os dados utilizados nos estudos do IBGE e do IPEA, da mesma forma em que não existe uma cultura acadêmica de compartilhar a descrição das metodologias utilizadas nas pesquisas, evidenciando que temos um problema de ‘divulgação de saberes especializados’ desde a produção das pesquisas, mesmo quando são públicas e sobre dados sociais públicos.

O desafio de ‘deter saberes’ é por vezes acoplado ao desafio de ‘proteção do banco de dados’ por políticas governamentais que proíbem a circulação de dados públicos em estratégias que vão desde a ‘proteção de relatórios’ até o descarte de documentos após o fim de um mandato governamental, as informações das empresas privadas então sequer podem ser ‘pedidos para acesso’. Assim, no Brasil, e em vários outros países, as áreas sociais mais frágeis e de difícil acesso são a segurança pública, a saúde e a propriedade intelectual, em que não só ‘falta

acesso ao armazenamento de dados', mas temos problemas com a proteção de informações que evidenciam problemas como violência policial, racismo institucional e negligência.

Diante desses desafios apontados, a partir de um trabalho puramente descritivo, é possível, visualizando as dificuldades de acesso e avaliação dos objetos da sociologia digital, pontuar rumos necessários para superar barreiras de formação, cultura, e economia. Quando Dominique Vinck (2013) propõe a sociologia digital como 'necessária a América Latina' ele faz menção a parte desses desafios apontando a necessidade de revertê-los para não cair em processos de controle social que reforçam práticas imperialistas e coloniais do poder-saber interferindo diretamente no patrimônio cultural e na construção de uma rede de humanização digital e globalizada. Mediante esses questionamentos, percebemos que por mais interessante e renovadora que seja a sociologia digital, sem revermos esses desafios, que começam desde o uso de metodologias quantitativas nas ciências sociais, no desenvolvimento de métodos e técnicas de pesquisa, a sociologia digital tende a ficar tão estacionada e restrita quanto os métodos quantitativos têm sido, um problema que afetará pesquisas sociais, promoção de políticas públicas e o próprio desenvolvimento da ciência e da tecnologia como 'saberes para detenção de poder'.

Referências

BAKSHY, Eytan; MESSING, Solomon; ADAMIC Lada. Exposure to ideologically diverse news and opinion on Facebook. *Science Express*, v. 1126, n. 2, p. 1- 10, 2015.

CASTRO-GÓMEZ, Santiago. Ciências sociais, violência epistêmica e o problema da "invenção do outro". In: LANDER, Edgardo (Org.). *A colonialidade do saber: eurocentrismo e ciências sociais. Perspectivas latino-americanas*. Buenos Aires: CLACSO, 2005. p. 169-186.

DOURISH, Paul. Algorithms and others: algorithmic culture in context. *Big Data & Society*, v. 3, n.2, p. 1-11, 2016.

FLICK, Uwe. *Introdução à metodologia de pesquisa*. Trad. de Magda Lopes: Porto Alegre: Penso, 2013.

FOWLER, Floyd J. *Pesquisa de levantamento*. Trad. de Rafael Padilha Ferreira:

Porto Alegre: Penso, 2011.

HIRATA, Helena. Gênero, classe e raça: interseccionalidades e consubstancialidades das relações sociais. *Tempo Social*, São Paulo, v. 26, n. 1, p. 61-73, 2015.

IBGE, Instituto Brasileiro de Geografia e Estatística. *Histórico*. Disponível em: <<http://www.ibge.gov.br/>>. Acesso em: 30 ago. 2016.

IPEA, Instituto de Pesquisa e Estatística Aplicada. *Cinco décadas de história*. Disponível em: <<http://www.ipea.gov.br/>>. Acesso em: 30 ago. 2016.

KERGOAT, Daniele. Consubstancialidade das relações sociais. *Novos Estudos*, n. 84, p. 93-103, 2010.

KERLINGER, Fred N. *Metodologia da pesquisa em ciências sociais*. São Paulo: EDUSP, 1980.

KUHN, Thomas. *A estrutura das revoluções científicas*. 6. ed. São Paulo: Perspectiva, 2001.

LINCOLN, Yvonna S.; GUBA, Egon G. Controvérsias paradigmáticas, contradições e confluências emergentes. In: DENZIN, Norman K.; LINCOLN, Yvonna S. *O planejamento da pesquisa qualitativa*. Porto Alegre: Artmed, 2006. p. 169-192.

MACFARLAND, Daniel A.; LEWIS, Kevin; GOLDBERG, Amir. Sociology in the Era of Big Data: The Ascent of Forensic Social Science. *American Sociologist*, New York, v. 47, n. 1, 2015.

MORAES, Fabiana. *No país do racismo institucional: dez anos de ações do GT Racismo no MPPE*. Recife: Procuradoria Geral de Justiça, 2013.

MUNIZ, Jerônimo Oliveira. Sobre o uso da variável raça-cor em estudos quantitativos. *Revista Sociologia Política*, Curitiba, v. 18, n. 36, p. 277-291, 2010.

MÜTZEL, Sophie. Facing Big Data: Making sociology relevant. *Big Data & Society*, v. 2, n. 2, 2015.

NASCIMENTO, Leonardo Fernandes. A Sociologia Digital: um desafio para o século XXI. *Sociologias*, Porto Alegre, 2016.

NOIRET, Serge. História Pública Digital. *Liinc em Revista*, Rio de Janeiro, v. 11, n. 1, p. 28-51, 2015.

ORTIZ, Renato. As Ciências Sociais e a Cultura. *Tempo Social*, v. 14, n. 1, p. 19-32, 2002.

PERELMAN, Chaïm; OLBRECHTS-TYTECA, Lucie. *Tratado da argumentação: a nova retórica*. São Paulo: Martins Fontes, 2002.

PETRUCCELLI, José Luis. Autoidentificação, identidade étnico-racial e

heteroclassificação. In: (Org.) PETRUCCELLI, José Luis; SABOIA, Ana Lúcia. Características Étnico-raciais da população, classificação e identidades. Rio de Janeiro: IBGE, 2013.

PIMENTA, Ricardo Medeiros. As rugosidades do ciberespaço: um contributo teórico aos estudos dos *web* espaços informacionais. *Informação & Sociedade: Estudos*, João Pessoa, v. 26, n. 2, 2016.

QUIJANO, Aníbal. Colonialidade do poder, eurocentrismo e América Latina. In: LANDER, Edgar (Org.). *A colonialidade do saber, eurocentrismo e ciências sociais. Perspectivas latino-americanas*. Buenos Aires: CLACSO, 2005. p. 227-278.

RAMOS, Marília Patta. Métodos quantitativos e pesquisa em ciências sociais: lógica e utilidade do uso da quantificação nas explicações dos fenômenos sociais. *Mediações*, Londrina, v. 18, n. 1, 2013.

REA, Louis M. e PARKER, Richard A. *Metodologia de pesquisa: do planejamento à execução*. Trad. de Nivaldo Montingelli Jr. São Paulo: Pioneira, 2000.

SARDENBERG, Cecília Maria Bacellar. Migrações Perigosas: As (des)venturas semânticas do conceito de gênero nos projetos e políticas para mulheres no Brasil. In: GONÇALVES, Eliane; VANNUCHI, Maria Lúcia; TRÓPIA, Patrícia Vieira; TOSTA, Tania Ludmila Dias (Org.). *Iguais? Gênero, trabalho e lutas sociais*. Goiânia: Editora da PUC Goiás, 2014. p. 19-38.

SPIVAK, Gayatri Chakravorty. La política de traducción. In: BARRET, Michèle; PHILLIPS, Anne (Org.) *Desestabilizar La Teoría*, Debates Feministas contemporáneos. Paídos Mexicana: México, 2002. p. 189-211.

STEPHAN, Beatriz González. Economías fundacionales. Diseño del cuerpo ciudadano. In: _____ (Comp.). *Cultura y tercer Mundo, 2. Nuevas identidades y ciudadanías*. Caracas: Nueva Sociedad. 1996. p. 17-47.

TOULMIN, Stephen. O lay-out de argumentos. In: _____. *Os usos do argumento*. 2. ed. São Paulo: Martins Fontes, 2006, p. 135-208.

VINCK, Dominique. Las culturas y humanidades digitales como nuevo desafío para el desarrollo de la ciencia y la tecnología en América Latina. *Universitas Humanística*, Bogotá, n. 76, p. 51-72, 2013.

WERNECK, Jurema (org.). *Racismo Institucional: uma abordagem conceitual*. Rio de Janeiro: Gelédes, 2013.

ZIEWITZ, Malte. Governing Algorithms: Myth, Mess, and Methods. *Science, Technology, & Human Values*, v. 41, n. 1, 2015.