

Estudando a cultura através de *selfies*: implicações do aspecto digital dos dados em sua análise¹²

Sandra Portella MONTARDO³

Eduardo PELLANDA⁴

André PASE⁵

Eduardo VELHO⁶

Universidade Feevale e Pontifícia Universidade Católica do Rio Grande do Sul

Resumo

Este artigo visa identificar as implicações que o aspecto digital constitutivo de *selfies* publicadas no Instagram podem ter sobre sua análise enquanto fenômeno cultural. A partir de diversos autores que abordam a importância de uma postura crítica na pesquisa de Humanidades e de Ciências Sociais em relação ao uso de dados digitais, descreveremos o processo de coleta, extração e classificação de *selfies* no Instagram por um programa elaborado no âmbito deste projeto de pesquisa. Concluímos que as *hashtags* constitutivas dos *selfies* exigem investigação mais aprofundada, por se referirem à questão da visibilidade, e que a submissão do *corpus* a outras formas de processamento de dados é pertinente para que possamos identificar com mais ênfase as particularidades do Instagram como plataforma em que os *selfies* foram coletados.

Palavras-chave: *Selfies*; Dados digitais; Metodologia; Instagram.

Introdução

O projeto no qual se insere este artigo tem por objetivo geral propor a base teórica da intersecção entre os Estudos de Performance e os Estudos de *Software* via a investigação das *selfies* como expressão de performances mediadas no site de rede social Instagram. Entre os objetivos específicos, está previsto identificar aspectos técnicos envolvidos no ato de se autofotografar e de compartilhar as *selfies* no Instagram, além de mapear usos e apropriações de usuários que postam *selfies* neste site

¹ Trabalho apresentado no GP Comunicação e Cultura Digital do XVII Encontro dos Grupos de Pesquisa em Comunicação, evento componente do 40º Congresso Brasileiro de Ciências da Comunicação.

² Artigo desenvolvido no âmbito do Projeto “Estudos de Performance como Estudos de Software: *selfies* como performance mediadas” (Edital Universal 01/16, Processo: 4203420311/2016-5).

³ Professora e Pesquisadora do PPG em Processos e Manifestações Culturais, do PPG em Diversidade Cultural e Inclusão Social e do Mestrado em Indústria Criativa, na Universidade Feevale, email: sandramontardo@feevale.br

⁴ Professor e Pesquisador do Programa de Pós-Graduação em Comunicação Social (PPGCom) e do Laboratório de Pesquisa em Mobilidade e Convergência Midiática (Ubilab) da PUCRS, email: eduardo.pellanda@pucrs

⁵ Professor e Pesquisador do Programa de Pós-Graduação em Comunicação Social (PPGCom) e do Laboratório de Pesquisa em Mobilidade e Convergência Midiática (Ubilab) da PUCRS, email: afpase@pucrs.br

⁶ Tecnólogo em Sistemas de Informática, Mestrando do PPG em Diversidade Cultural e Inclusão Social, da Universidade Feevale, e-mail: velho@feevale.br

de rede social, para que se possa, finalmente, conceituar performances mediadas, nos termos propostos.

Como pano de fundo para essa investigação, parte-se da suposição de que os Estudos de Performance carecem de tensionamento em relação à questão da tecnologia em sua abordagem, devido a uma nova configuração social implicada tecnicamente. De forma mais explícita, pretende-se investigar por quais meios os componentes técnicos que tornam o fenômeno dos *selfies* possível conformam aspectos relacionados a processos de subjetivação contemporâneos em termos de performances mediadas.

Para tanto, recorre-se, teoricamente, aos Estudos de Performance, que, a partir da proposição de seu fundador e maior representante, Richard Schechner (2002), dispõe a tecnologia como um tipo de performance entre outras, quando, na verdade, ela parece atravessar e constituir todas as demais, bem como à abordagem conhecida como Estudos de *Software*, que tem como premissa a centralidade da influência dos fatores computacionais em práticas culturais por meio de apropriações. Em relação aos Estudos de *Software*, trata-se de uma abordagem que privilegia a inserção dos *softwares* em tarefas cotidianas e que prevê consequências sociais, culturais e subjetivas provenientes de seus usos. Portanto, pode-se dizer que o foco dos Estudos de *Software* está nos usos e apropriações que os sujeitos fazem dessas várias “camadas de *software*” (Manovich, 2013) e não da simples descrição de funcionamento desses dispositivos. Nesse sentido, de acordo com Fuller (2008), mais do que apenas código, os *softwares* devem ser entendidos pelo seu valor simbólico de intervenção nas práticas culturais. De forma pragmática, Fuller (2008) define os Estudos de *Software* como um conjunto de estudos breves sob forma de textos críticos, expositivos e especulativos sobre objetos digitais particulares, linguagens e estruturas lógicas.

Diretamente associado aos nossos interesses de pesquisa, percebe-se, na última década, a emergência da discussão sobre a especificidade dos dados digitais na pesquisa de Humanidades. Nesse sentido, o livro organizado por Schäfer e Es, *The datafied society: studying culture through data* (2017), cujo subtítulo inspirou o título deste artigo, oferece um panorama de práticas de pesquisa com dados digitais, em especial, voltado para a pesquisa de mídia. Os organizadores (2017) destacam que ver nos dados digitais a possibilidade de entender o comportamento humano, requer saber como eles são gerados, coletados, selecionados e analisados (Langlois et al, 2015), destacando

ainda que esses processos afetam o modo como o conhecimento é produzido (Weinberger, 2014; Meyer; Schroeder, 2015).

Mais relacionado com o primeiro objetivo específico denominado acima, mas não restrito a ele, este artigo parte das seguintes perguntas: que limites e possibilidades de pesquisa os *selfies* publicadas no Instagram oferecem quando captadas como dados digitais? Que questões são postas pelo aspecto digital desses dados e que interferem naquilo que possamos compreender sobre os *selfies* enquanto fenômeno cultural?

Frente a isso, o objetivo deste artigo é identificar as implicações dos *selfies* publicadas no Instagram enquanto dados digitais a serem captados, processados e analisados no âmbito de uma pesquisa de Ciências Sociais Aplicadas. Em termos metodológicos, este estudo se vale de levantamento bibliográfico e de pesquisa documental.

1. Desafios do uso de dados digitais nas Ciências Humanas e Sociais

A introdução desse artigo evidencia o compromisso do projeto em pontuar as interferências digitais constitutivas de práticas culturais que incluem dispositivos digitais (ex: *selfies* postadas no Instagram). Gunthert (2015) define *selfie* como “imagem autofotográfica compartilhada” (GUNTHERT, 2015, p. 15) e estranha o fato de as mesmas terem se tornado o centro de debates de mídia apenas em 2013⁷, apesar de terem surgido no Japão já em 2000. Para o autor (2015), esse *delay* se deve à resistência a práticas culturais tornadas possíveis pelo que chama de ferramentas de conectividade, o que coloca as *selfies* como um sintoma dessa cultura conectada. Desta forma, a imagem capturada e difundida digitalmente também difere de registros próprios dos fotógrafos, seja diante de espelhos ou outras estratégias que permitissem tal registro.

A partir de uma abordagem estrutural, Artieri et al. (2015) propõem que *selfies* sejam “fotos instantâneas pessoais que a) são criadas com a consciência de pertencerem a um gênero específico de autorretratos; b) que combinam uma dimensão textual e icônica; c) são espalhadas nas redes sociais de indivíduos” (ARTIERE et al., 2015, p. 17). Os autores (2015) partem da abordagem de Goffman para compreender a prática das *selfies* em termos de efeitos de enquadramento, especificamente, a partir de dois fatores: a) a locação em que a foto foi tirada; b) as legendas e a *hashtag* que dispõem

⁷ Em 2013, *selfie* foi escolhida como a palavra do ano pela editora do Dicionário Oxford, um dos mais representativos do planeta. (BBC NEWS, 2013). Disponível em: Disponível em <http://www.bbc.com/news/uk-24992393>. Acesso em 8 de julho de 2017.

esse gênero de foto no “fluxo das imagens de mídias sociais” (ARTIERE et al., 2015, p. 17).

Por conta da natureza de nosso objeto de pesquisa, é válido acompanhar o que vem sendo produzido no sentido de destacar a especificidade dos métodos digitais e o que significa fazer pesquisa em Humanidades/Ciências Sociais Aplicadas com a utilização de dados digitais.

Em uma sociedade em que muitos aspectos da linguagem, discurso e cultura são transformados em dados para sua circulação/comunicação, processo coloquialmente chamado de “dataficação”, é fundamental investigar as condições e o contexto do qual eles emanam (Van Dijck, 2017). Na medida em que, atualmente e de forma legítima, pode-se acessar, entender e monitorar comportamentos humanos, Es e Schäfer (2017) destacam que a análise de dados promete uma forma objetiva de dominarmos a vida complexa e dinâmica em que vivemos, ainda que a tradução do social em dados requiera atenção sobre como esses dados foram gerados e por que meios eles dizem algo sobre alguma questão.

Em outras palavras, mais do que analisar fenômenos pelos dados a que temos acesso, as Humanidades devem desenvolver novas ferramentas conceituais para verificar seus impactos na sociedade. Nesse mesmo sentido, Ramsay (2003), citado pelos organizadores do livro, afirma que o propósito da pesquisa de dados pelas Humanidades não é apenas validação empírica e teste de hipóteses, mas desenvolvimento de questões e descoberta de *insights*.

Masson (2017) chama atenção para a questão interdisciplinar envolvida no uso de dados digitais pelos pesquisadores de Humanidades, destacando que quando fazemos uso desse tipo de dados, incorporamos as tradições epistêmicas que derivam deles. Nesse sentido, este autor se vale de Drucker (2012) para advertir sobre o perigo de se utilizar as imagens geradas por visualizadores de dados como se elas fossem a própria realidade representada por eles, e sobre a importância de se atribuir significado a dados gerados digitalmente.

Ainda quanto a isso, Masson (2017) retoma Drucker (2014) para afirmar que a tradução de informação em forma gráfica promove simplicidade e legibilidade que, na verdade, escondem aspectos originais da construção dos dados. Drucker (2014) também salienta que dados são sempre pré-constituídos de acordo com os parâmetros que guiam

a sua seleção. De acordo com esse ponto de vista, dados (*data*) não são dados (*given*), mas *capta*, tomados (*taken*), portanto, construídos como uma interpretação do mundo.

Manovich (2017) fala sobre a Analítica Cultural como uma perspectiva de pesquisa que combina Computação Social e Humanidades Digitais. Segundo o autor (2017), as Humanidades Digitais analisam artefatos históricos criados por escritores, artistas e músicos, que, em função de *copyright*, resultam em amostras pequenas. Já a Computação Social parte de amostras, frequentemente, de conteúdos gerados pelos usuários e das interações empreendidas entre eles, milhares de vezes maiores do que as estudadas pelas Humanidades Digitais, ainda que, apenas, detectem padrões, sem encontrar suas especificidades locais. Frente a isso, o objetivo da Analítica Cultural seria combinar tanto a análise de grandes volumes de dados com um viés local. São exemplos de projetos baseados na Analítica Cultural e coordenados por Manovich as iniciativas Selfiecity⁸ e OnBroadway⁹.

Segundo Rieder e Röhle (2017), frente à proliferação de termos como Humanidades Digitais (*Digital Humanities*), Analítica Cultural, métodos digitais ou Web Science, Berry (2011) vislumbrou o que chamou de virada computacional, que diz respeito, não apenas, ao aumento de uso de dados quantitativos na pesquisa social e cultural, mas que designa a manipulação de conjuntos de dados maiores, envolvendo técnicas de visualização avançadas e processamento de dados, às vezes, confusos. Com o intuito de lidar com as transformações em curso que desafiam práticas epistemológicas e paradigmas estabelecidos, Rieder e Röhle (2017) enumeram 5 desafios que a virada computacional impõe aos pesquisadores de Ciências Humanas e Sociais:

- a) A isca da objetividade: relativizar a premissa de que mais dados e processamento de dados representam objetividade, de modo que um olhar interpretativo e associado ao contexto social de sua geração deve ser levado em conta.
- b) O poder da evidência visual: prevê a combinação de crítica e interpretação para a identificação de *insights* a partir de formas visuais produzidas por computador.

⁸ Selfiecity é um projeto que analisou 3200 *selfies* capturados e marcados com informações relativas ao seu posicionamento geográfico em Bangkok, Berlim, Moscou, Nova York e São Paulo, cujo material foi publicado no website <http://selfiecity.net>.

⁹ On Broadway é uma instalação derivada do projeto Selfiecity que explora a Avenida Broadway, uma das principais de Nova York, usando 40 milhões de imagens geradas por usuários e comparando seus dados com outros como uso de táxis e renda média dos moradores dos locais. O seu conteúdo está disponível em <http://www.on-broadway.nyc>.

-
- c) *Black boxing*: consiste em saber como o método funciona e criticá-lo.
 - d) Perturbações Institucionais – diz respeito à interação de pesquisadores de áreas diferentes em um mesmo projeto, de modo que cada área esteja devidamente representada e defendida pelos pesquisadores.
 - e) Em busca do Universalismo – refere-se à combinação de diferentes configurações metodológicas, tanto digitais quanto não digitais.

A partir de Berry (2011), Es, Coombs e Boeschoten (2017) afirmam que, embora a análise de dados não constitua uma prática nova, a quantidade de dados disponíveis atualmente consiste em novos recursos para explorar práticas sociais e relacionamentos, tornando novas análises culturais possíveis. No entanto, também esses autores (2017) fazem lembrar que os dados não constituem um fenômeno natural, uma vez que existem em um contexto social particular. Frente a isso, antes de proporem uma série de questionamentos que devem embasar a coleta, a seleção e análise dos dados, Es, Coombs e Boeschoten (2017) destacam o que consideram como erros de interpretação referentes a esse tema. Um primeiro seria o pressuposto de que dados fariam por si só, o que tornaria qualquer método científico obsoleto (Anderson, 2008). Quanto a isso, os autores (2017) retomam Kitchin (2014b) e Couldry (2016), que defendem que nenhuma produção de conhecimento é livre de interpretação. Associado a esse primeiro mal-entendido, emerge um segundo, a partir do qual costuma-se dizer que grandes quantidades de dados promovem a detecção de padrões, quando, segundo esses autores (2017), a combinação de análise qualitativa é necessária para fazer a parte quantitativa dos dados fazer sentido de acordo com o que se pretende investigar. De forma resumida, Es, Coombs e Boeschoten sinalizam alguns pontos que devem ser levados em conta na análise de dados digitais na pesquisa social e de humanidades:

- De onde os dados vem;
- Quem produziu os dados e com qual propósito;
- Que dados são selecionados e como eles se relacionam com o conjunto de dados maior;
- Que ferramentas foram usadas para coleta e análise;
- Por que certos dados e métricas foram usados para a pesquisa; (ES, COOMBS, BOESCHOTEN, 2017, p. 178).

No intuito de diferenciar métodos de pesquisa que surgiram com a web (sistemas de recomendação, folksonomia) daqueles que foram adaptados para ela (diretórios, observação participante e não participante, entrevistas on-line), Rogers (2016) propõe

um programa chamado “Digital Methods”, prática de pesquisa que “aprenda com os métodos e dispositivos on-line, readapte-os e busque fundamentar afirmações sobre a mudança cultural em dados web, introduzindo o termo “fundamentação em dados online” (*online groundedness*).” (ROGERS, 2016, p. 2). O autor (2016) argumenta, entre outros pontos, que o estudo de usuários (amadores, fãs, produzidos) desloca a atenção para longe dos dados digitais produzidos por eles.

“A questão não é mais o quanto a sociedade e cultura estão on-line, mas sim como diagnosticar mudanças culturais e condições sociais por meio da internet” (ROGERS, 2016, p.5), de modo que a Internet seja tomada não como objeto de estudo, mas como fonte de estudo. Um exemplo disso é o Google Flu Trends, projeto lançado em 2008, e que antecipa e localiza geograficamente possibilidades de surto de *influenza*, pela quantificação de buscas realizadas, dados que podem ser utilizados em conjunto com outras mensurações existentes (número de hospitalizações, de pacientes, relatórios de laboratório, etc.). (ROGERS, 2016). Outra possibilidade que toma a internet como fonte de pesquisa aplicada a sites de redes sociais, em projeto desenvolvido por Rogers (2016), parte das informações de gostos e preferências declaradas nesses ambientes pelos usuários para propor uma nova forma de se definir um público, processo que culminou com o seu conceito de pós-demografia¹⁰.

A citação a seguir amplia a questão de como estudar a cultura e sociedade e fundamentar as descobertas com a internet.

Coletar e analisar esses dados para a pesquisa social e cultural requer não apenas uma nova perspectiva sobre a internet, mas também novos métodos para fundamentar as descobertas. Fundamentar as reivindicações nas especificidades do on-line é uma mudança maior no propósito da pesquisa em internet, na medida em que não se trata tanto de pesquisar a internet e seus usuários, mas sim, estudar a cultura e a sociedade com a internet. (ROGERS, 2016, p. 30)

Frente ao que foi posto, pergunta-se: o que se torna possível de ser feito culturalmente/socialmente pelos usuários, por meio de e apesar das especificidades técnicas dos *selfies*, envolvendo o Instagram e os dispositivos que os capturam? Acreditamos que podemos chegar a essa resposta de diferentes maneiras. A escolhida nesse artigo é investigar as implicações de coleta, extração e classificação do *corpus* por

¹⁰ Pós-demografia consiste na composição de públicos a partir de suas preferências declaradas em sites de redes sociais, e que vão além dos dados demográficos para tanto (Rogers, 2016).

meio da proposta de construção de uma rotina automatizada para análise de *selfies* publicadas na plataforma Instagram.

2. Processo de coleta, extração e classificação do *corpus*

Esta seção do artigo refere-se ao item sobre ferramentas utilizadas para coleta de dados, do roteiro proposto por Es, Coombs e Boeschoten (2017). Sabe-se que o processo de coleta de dados busca compor o *corpus* do estudo. Cada unidade do *corpus* é composta por uma publicação, que possui uma imagem com seus respectivos parâmetros (*hashtags*, número de curtidas e legenda, por exemplo). Estes dados foram retirados do site de rede social Instagram, seguindo as diretrizes apresentadas no regulamento do próprio site (FACEBOOK, 2013)¹¹. Desta forma, todo o processo de coleta de dados ocorreu de forma manual, sem qualquer tipo de automação; pois, como descreve o item 10 das “Condições gerais” dos Termos de Uso do Instagram,

Nós proibimos o rastreamento, *scraping*, armazenamento em *cache* ou, de qualquer outro modo, o acesso a qualquer conteúdo no Serviço através de meios automatizados, incluindo, entre outras coisas, perfis e fotos de usuários (exceto como resultado de protocolos ou tecnologias de mecanismos de busca padrão com a autorização expressa do Instagram). (FACEBOOK, 2013).

Neste sentido, nenhum *web crawler*¹², *web scrapper*¹³ ou qualquer outro tipo de processo automatizado foi utilizado para realizar esta coleta de dados. Mesmo a *API*¹⁴ do Instagram não pode ser utilizada. Esta *API* possui como principal objetivo oferecer uma interface de comunicação entre uma aplicação de *software* cliente e o serviço Instagram; mas não pode ser utilizada para esta pesquisa devido às restrições das regras de negócio desta interface. Isto é vetado pelos termos: 8, 14, 16 e 17 do vigente contrato.

Existem ferramentas como o *instagram-php-scrapper*¹⁵, que permitem capturar automaticamente um conjunto de dados do Instagram, porém não foram utilizadas pois

¹¹ O Instagram foi adquirido pelo Facebook em 2012 (RUSLI, 2012), desta forma seus termos de serviço estão creditados para a empresa que detém o comando sobre a outra.

¹² *Software* que recursivamente acessa *links* de um site e faz *download* de todas as páginas acessadas. *Web crawlers* podem ser orientados por regras, como por exemplo, podem ser configurados para somente seguir *links* com o sufixo “.html” ou “.php”.

¹³ *Software* desenhado para extrair informações de uma determinada página da *web*. Esta filtragem serve geralmente para indexar informação para que então possa ser armazenada em uma base de dados.

¹⁴ *Application Programming Interface*. Uma interface desenhada para programadores acessarem com mais facilidade determinados componentes de *software*.

¹⁵ *Scrapper* não-oficial para Instagram. Disponível em: <<https://github.com/postaddictme/instagram-php-scrapers>>. Acesso em: 19 de jun. 2017.

as diretrizes do Instagram impedem seu uso. Além disso, eventuais atualizações da API culminam por invalidar serviços fornecidos por terceiros (Fantoni, 2017).

Como alternativa às formas automatizadas de coleta de dados, foi utilizado um método manual de acesso ao Instagram. O objetivo deste método é salvar em *off-line* os dados do Instagram para então extrair de forma automatizada somente o que é relevante no contexto deste estudo. Isto também evita que a estratégia da pesquisa dependa de um serviço alheio ao seu controle e seja fragilizada por uma eventual mudança na rede.

O processo manual de coleta de dados consiste em acessar uma publicação do Instagram em sua visão modal [onde são listadas as *hashtags*, comentários, curtidas e outros parâmetros associados à publicação] e salvar no computador (em *off-line*) o conteúdo *HTML* respectivo ao endereço acessado. Este conteúdo *HTML* carrega informações acerca da publicação respectiva, que podem ser extraídas programaticamente utilizando um *software* do tipo *scraper*. Porém, não são todas as informações associadas às publicações que estão neste conteúdo *HTML*, visto que algumas destas informações são carregadas dinamicamente ao acessar as páginas com um *web browser*.

A escolha de salvar o conteúdo *HTML* das publicações deve-se ao fato de que, com estes arquivos salvos em *off-line*, torna-se possível automatizar parte do processo de coleta de dados sem ferir os Termos de Contrato do Instagram. Esta ação é similar a um processo de registro das informações observadas por um interagente comum que acessa tal foto (ou tal endereço, se o acesso é realizado pela versão *web* da rede). As diretrizes do Instagram proíbem somente o *acesso* automatizado ao serviço (FACEBOOK, 2013); se este acesso é realizado de forma manual, os dados salvos a partir destes acessos podem ser operados com processos automatizados em *off-line*, sem realizar qualquer tipo de acesso ao serviço Instagram.



Fonte: Elaborada pelos autores

A parte automatizada da coleta de dados consiste em extrair e organizar as informações existentes nos arquivos de conteúdo *HTML* (informação não-filtrada) obtidos no Instagram. Os dados extraídos são: legenda, quantidade de comentários, *hashtags*, quantidade de *likes*, *link* de acesso à imagem da publicação e, o mais importante, a imagem de fato. O material deve ser capturado para posterior armazenamento em arquivos *JSON* (informação filtrada), acessados programaticamente para a realização das tarefas de análise.

Desta forma, os arquivos da pesquisa possuem somente as informações que são relevantes para o estudo. Cada arquivo *HTML* ordenado pelo script *JSON* reproduz uma determinada publicação do Instagram, porém adaptada para os padrões do projeto. Além de evitar que uma eventual imagem importante seja deletada e, conseqüentemente, perdida, permite organizar o conteúdo dentro de uma estrutura adaptada para o projeto, sem a necessidade de constantes procuras e visitas através da versão web do serviço.

O *corpus* do estudo é composto somente por publicações em que a imagem associada é um *selfie*. As publicações foram acessadas manualmente e avaliadas caso por caso, salvando aquelas em que o sujeito que maneja o dispositivo de captura está presente na imagem capturada. É necessário lembrar que um *selfie* pode exibir outras pessoas ou objetos além do indivíduo que realizou a captura da imagem, porém este deve estar presente na foto.

As características mais notáveis do *selfie* são: a presença de um “pau de *selfie*”¹⁶ na imagem; presença de algum sujeito com os braços levantados em direção ao dispositivo de captura de forma que indique que somente algum indivíduo presente na fotografia poderia ser o responsável por operar o dispositivo de captura; e, por fim, fotos onde o dispositivo de captura aparece sendo operado por um sujeito (como fotos de espelho ou capturadas a partir de um *drone*) presente no registro.

Para obter as publicações que compõem o *corpus*, a função de busca do Instagram foi utilizada. Os parâmetros utilizados no processo de filtragem foram *hashtags* e localizações. As *hashtags* utilizadas levam em sua semântica tópicos relativos à temática dos *selfies*, como *gopro*¹⁷, *nofilter*, *selfiestick* e *wefie*, por exemplo.

¹⁶ Espécie de bastão usado para estender a área capturada na imagem, formado por uma base para prender um *smartphone* ou câmera em uma extremidade e um botão para disparo da foto na outra.

¹⁷ Em virtude do seu tamanho reduzido e da utilização sobretudo para registro de atividade esportiva, as câmeras da empresa GoPro são muito utilizadas para a captura de *selfies*. Apesar de não existir um registro inicial do uso do “pau de *selfie*”, a empresa foi uma das primeiras a oferecer o extensor para os seus consumidores.

As localizações utilizadas são referentes a pontos turísticos brasileiros. Ao inserir os parâmetros no campo de busca, as primeiras 50 publicações exibidas, que levavam uma imagem classificada como *selfie* foram selecionadas para o estudo. Em alguns casos a busca exibiu um número inferior a 50 *selfies*. Nesses casos, uma menor quantidade de publicações foi selecionada para o *corpus* do estudo.

Para facilitar a visualização dos dados, as imagens foram classificadas por *hashtags*, possibilitando que somente publicações que possuem uma determinada *hashtag* sejam filtradas. As imagens classificadas dentro de cada *hashtag* foram unidas em uma *grid* para facilitar a visão em um panorama geral. Cada uma destas imagens possui uma legenda com os parâmetros existentes no arquivo *JSON* da publicação respectiva.

Após a finalização da coleta de dados, extração e classificação do *corpus*, a base de dados foi verificada por publicações repetidas ou com algum tipo de informação corrompida. As publicações que encaixam nestes casos foram removidas do *corpus*. Por fim, restaram 1991 publicações com um somatório de 4430 *hashtags* únicas.

Este grupo, portanto, foi selecionado para uma posterior análise. Em virtude do volume do material e da sua análise realizada de maneira que observa caso por caso, a equipe, atualmente, está analisando as imagens. Este material será a base do segundo passo da pesquisa e desenvolvido para futura apresentação. Cabe recordar que esta combinação de métodos automatizados com separação manual é inspirada no projeto Selfiecity (2014) descrito anteriormente e que combina separação automática com análise segmentada das publicações.

5. Observações Preliminares e Limites da Pesquisa

Pode-se dizer que a análise de *selfies* no Instagram envolve, ao mesmo tempo, uma compreensão macro sobre as dinâmicas do capitalismo contemporâneo, sobre distribuição de informação orientada através de algoritmos, uso de dispositivos de comunicação ubíquos, análise do conteúdo que constitui tais redes (Fantoni, 2017) e, finalmente, o propósito dos usos praticados pelos usuários dessa plataforma.

A partir do roteiro sugerido por Es, Coombs e Boeschoten (2017), os *selfies* que compõem o *corpus* dessa pesquisa vem do Instagram, que foi escolhido como fonte desse tipo de imagem pelo fato de ter sido o primeiro site de rede social projetado para dispositivos móveis, aspecto que incide diretamente na captura desse tipo de imagem e

na facilidade de seu compartilhamento. Além disso, estudo de Wu et al. (2014) a partir dos usos de 40 mil usuários de 179 países, destacou *selfies* como o tipo de foto mais postado no Instagram entre 8 tipos identificados (*selfies*, amigos, moda, comida, *gadgets*, atividades, animais de estimação e fotos com texto).

A questão da mobilidade inerente ao Instagram diferencia este site de rede social de outros que não eram focados neste aspecto, inicialmente, pela questão da geolocalização. Por exemplo, o conteúdo de redes como o Orkut estava vinculado a uma publicação online, com eventuais referências para locais apenas relativas ao seu conteúdo. Essa dinâmica era semelhante, no início, do Facebook, porém foi transformada após o surgimento de híbridos de rede e serviço como Foursquare (e sua futura cisão para o Swarm) e GoWalla. Além de utilizar a geolocalização como fator principal, com publicações relativas a um local – principalmente dicas de restaurantes e estabelecimentos –, esses serviços encampavam a mobilidade ao centralizar seu foco de uso nos usuários de *smartphones*. Desta forma, a transformação do site de rede social em um aplicativo acompanhado dos detalhes geográficos permitiu usar as redes para os mais diversos fins. Isto é observado na comunhão entre rede social digital, serviço e aplicativo observada em serviços de encontros como Tinder e Grindr.

Refletir sobre “Que dados são selecionados e como eles se relacionam com o conjunto de dados maior” (Es, Coombs e Boeshoten, 2017, p. 178) remete à já mencionada geolocalização e a *hashtags*. No caso deste estudo, a geolocalização constituiu o parâmetro inicial para seleção de *selfies*, a partir de pontos turísticos de vários locais do Brasil, o que, ao mesmo tempo, assegura abrangência de seleção de *selfies* postados por usuários em diferentes regiões geográficas na seleção do *corpus*, mas que, também, sugere a predominância de fotos ao ar livre, em locais de grande circulação e com acesso à rede.

Como se sabe, o Instagram compreende a imagem como uma unidade importante para o seu espaço de visualização e, sobretudo, do seu aparelho – e que configura a sua plataforma. Como observa Fantoni (2007), em relação à forma como jovens de Porto Alegre realizam atos de autorrepresentação na plataforma, o uso do aparelho como uma câmera que está conectada não apenas representa imagens relevantes para o seu autor, mas permitem que elas fomentem pontos de discussão sobre seu conteúdo e também sirvam como uma espécie de marcadores de discussões e/ou locais, quando vinculados por *hashtags* e dados de localização. Além disso, a fruição do

seu conteúdo é fácil em virtude das propriedades dos aparelhos, sobretudo quanto ao acesso das fotos via o ato de rolar com o dedo para obter mais publicações e apertar duas vezes seguidas para “favoritar” uma imagem ou vídeo.

Em virtude de ser uma rede com objetivos comerciais, o seu uso é voltado para a geração de capital para a empresa, seja ele social ou financeiro de fato, o que resulta em alterações nas dinâmicas da rede. Por outras palavras, o Instagram, assim como outras plataformas, utiliza os dados de socialização de seus usuários para gerar anúncios que, a princípio, sejam do interesse desses usuários e que, por conta disso, convertam-se em vendas para seus anunciantes. Para tanto, é conveniente para o Instagram que os usuários sejam ativos na plataforma. Entre exemplos de adaptações nesse sentido, estão a adoção de novos recursos, como as *stories*, ou supressão de outros, como alterações na API, que impedem o uso de serviços como o IfThisThanThat¹⁸ para automatização de rotinas ou a redução de espaço na indicação da data da postagem, dados importantes para a compreensão da rede acabam, cada vez mais, restritos apenas para o Facebook, uma vez que são os recursos mais valiosos da plataforma e que despertam o interesse do mercado como um todo.

Isto pode ser observado na forma como algoritmos operam na distribuição de conteúdo. No entanto, o material que forma o *corpus* não está tão sujeito às dinâmicas do tempo, quanto à sua permanência e distribuição nas *timelines*. Ainda que fotos pessoais e as postadas por um veículo de comunicação obedeçam a diferentes dinâmicas temporais quando disponibilizadas no mesmo site de rede social, pelo fato de os *selfies* do *corpus* terem sido coletados manualmente, essa interferência do algoritmo é amenizada.

A análise proposta também permite compreender que uma rede é formada não apenas pelos seus atores, mas pelas suas representações visuais. A análise de comentários e *hashtags* é importante para criar contextos que vão além daqueles representados imagetivamente – ainda que estejam conectados a estes.

Além disso, serão observadas propriedades relativas aos aparelhos utilizados pelos interagentes. A seleção do *corpus* de pesquisa por *hashtags* como #drone e #gopro, por exemplo, são indicadores do uso desses dispositivos para a captação de *selfies*.

¹⁸ Serviço que permite a criação de diferentes rotinas automatizadas definidas pelo usuário, disponível online em <https://iftt.com>.

Considerações finais

Ainda que o fenômeno dos *selfies* não seja restrito ao site de rede social Instagram, uma vez que, segundo Gunthert, remontam ao ano 2000, arrisca-se que o desenvolvimento de variadas plataformas e de dispositivos de captura desse tipo de imagem parece ter atualizado esse fenômeno no que hoje se pode observar quanto a este tipo de postagem no Instagram.

Pelo que notamos, a partir dos critérios de seleção do *corpus*, as *hashtags* figuram como ponto de destaque a ser investigado, uma vez que, remetem tanto à anti-figuratividade dos *selfies* (Di Foggia, 2015), ao lado da geolocalização, quanto à questão da importância da visibilidade para seus usuários (Fantoni, 2017). Esse último ponto é realçado quando notamos, por exemplo, a popularidade de geradores de *hashtags* para o Instagram¹⁹. Esse tipo de *insight* está associado à abordagem de Richards (2016), segundo o qual, Métodos Digitais dizem respeito a como podemos pensar a sociedade utilizando a Internet como fonte.

Submeter o *corpus* da pesquisa ao Google Fotos parece ser uma alternativa de observarmos as possibilidades de tratamento de dados digitais pela ótica de um outro grande player da Comunicação Digital mundial, o que pode trazer *insights* interessantes para o prosseguimento da pesquisa.

Finalmente, cabe ressaltar a importância dos resultados alcançados para subsidiar o segundo objetivo específico do projeto, qual seja, mapear usos e apropriações de pessoas que postam *selfies* por meio de questionários e de entrevistas em profundidade, atendendo à questão sobre quem produziu os dados e com qual propósito (Es, Coombs, Boeschoten, 2017).

REFERÊNCIAS

ARTIERE, G. B.; FARCI, M.; ZUROVAC, E.; OREFICE, M. The framing effect of selfies. Pictures beyond photography. In: #selfie. Imag(in)ing the self in digital media. Book of abstracts. 2015. Disponível em: <http://www.uni-marburg.de/fb09/medienwissenschaft/forschung/veranstaltungen/selfie_bilder-dateien/selfie_abstracts.pdf>. Acesso em: mar. 2017.

DI FOGGIA, G. About the anti-figurativeness of #selfie. In: #selfie. Imag(in)ing the self in digital media. Book of abstracts. 2015. Disponível em: <http://www.uni-marburg.de/fb09/medienwissenschaft/forschung/veranstaltungen/selfie_bilder-dateien/selfie_abstracts.pdf>. Acesso em: jul. 2017.

¹⁹ <http://www.techtudo.com.br/listas/noticia/2016/08/sete-apps-para-encontrar-melhores-hashtags-para-usar-no-instagram.html>. Acesso em: 13 jul. 2017.

ES, K. V.; COOMBS, N. L.; BOESCHTEN, T. Towards a Reflexive Digital Data Analysis. In: SCHÄFER, M. T. & ES, K. V. (Ed.) **The datafied society**. Studying Culture through data. Amsterdam: Amsterdam University Press, 2017.

FANTONI, A. Autorrepresentação de Adolescentes Porto-Alegrenses no Instagram. 2017. 175 f. **Dissertação** (Mestrado em Comunicação Social) – Faculdade de Comunicação Social, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2017.

FULLER, M. (Org.) **Software Studies**: a lexicon. Cambridge, MA: MIT Press, 2008.

GUNTHER, A. The selfie: disease and emblem of the connected image. In: #selfie. Imag(in)ing the self in digital media. Book of abstracts. 2015. Disponível em: http://www.unimarburg.de/fb09/medienwissenschaft/forschung/veranstaltungen/selfie_bilder-dateien/selfie_abstracts.pdf. Acesso em: 26 mar. 2015.

JSON.ORG. **Introdução ao JSON**. Disponível em <http://www.json.org/json-pt.html>. Acesso em 8 de julho de 2017.

MANOVICH, L. Cultural Analytics, Social Computing and Digital Humanities In: SCHÄFER, M. T. & ES, K. V. (Ed.) **The datafied society**. Studying Culture through data. Amsterdam: Amsterdam University Press, 2017.

MANOVICH, L. **Software takes commande**. Extending the language of new media. Nova York, NY: Bloomsbury Academic, 2013.

MASSON, E. Humanistic Data Research. An Encounter between Epistemic Traditions. In: SCHÄFER, M. T. & ES, K. V. (Ed.) **The datafied society**. Studying Culture through data. Amsterdam: Amsterdam University Press, 2017.

ROGERS, R. O fim do virtual. Os métodos digitais. **Lumina**, v. 10, n. 3, dez. 2016. Disponível em: <<https://lumina.ufjf.emnuvens.com.br/lumina/article/view/644>>. Acesso em: jun. 2017.

SCHÄFER, M. T. & ES, K. V. (Ed.) **The datafied society**. Studying Culture through data. Amsterdam: Amsterdam University Press, 2017.

SCHECHNER, R. **Performance studies** – an introduction. NY: Routledge, 2002.

VAN DIJCK, J. 2017. Foreword. In: SCHÄFER, M. T. & ES, K. V. (Ed.) **The datafied society**. Studying Culture through data. Amsterdam: Amsterdam University Press, 2017.

WU, Y.; MANIKONDA, L.; KAMBHAMPATI, S. What we Instagram: a first analysis of Instagram photo content and user types. Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media, 2014, pp. 595-598. Disponível em: <https://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/view/8118>. Acesso em: 13 jul. 2017.