

Pesquisa baseada em Dados Sociais Digitais: mapeamento de ferramentas e táticas de coleta de dados no Intercom.

Tarcízio Silva (Brasil).¹

Resumo.

O trabalho apresenta análise bibliográfica de produção dos Grupos de Pesquisa Cibercultura e Conteúdos Digitais & Convergências Tecnológicas do Congresso da Intercom (Brasil) nos últimos dez anos. Nos debruçamos sobre os artigos que realizaram coleta de dados interacionais, conversacionais e opinativos em sites de redes sociais. Variáveis relacionadas a quantidade de dados, ferramentas de coleta, plataformas analisadas e apresentação dos resultados foram rastreadas à luz de questões referentes aos desafios e potencialidades destas modalidades de pesquisa.

Palavras-chave.

sites de redes sociais; coleta de dados; pesquisa empírica; Intercom.

Abstract.

The paper presents literature review of the production of Cyberculture and Digital Content & Technological Convergence Research Groups of Intercom Congress in the last ten years. We look back on the articles that performed gathering of interactional, conversational and opinion data on social networking sites. Variables related to the amount of data collected, platforms analyzed and presentation of results were screened in the light of issues related to the challenges and potentials of these research methods.

Keywords.

Social network sites; data collection; empiric research; Intercom.

Pesquisa Baseada em Dados Empíricos dos Sites de Redes Sociais.

A virada computacional na produção de conhecimento científico não é nada subestimada. Hoje é praticamente consenso que a ubiquidade dos computadores e acesso à internet em algumas sociedades gera um impacto profundo nas humanidades (Manovich, 2012), ciências sociais (Savage & Burrows, 2009) e na comunicação de modo particular (Scolari, 2009). As profundas transformações na sociedade, indústria, mercado, política, relações de trabalho, concepções de lazer e na vida cotidiana de modo geral agregam novos problemas, questões e possibilidades pertinentes aos dispositivos tecnológicos em si e suas *affordances*, ao papel do acesso e produção de dados pelos indivíduos e a própria gestão, armazenamento e redesignação destes dados por diferentes atores sociais.

Um recorte sobre métodos para as ciências sociais, usando como malha norteadora o trabalho clássico de Andrew Abbott (2004), pode ser utilizado para compreender o impacto do digital e comunicação digital na pesquisa em comunicação. Para dar conta dos diferentes métodos das ciências sociais, Abbott as descreve em diversos níveis. O primeiro nível se refere ao método de coleta de dados e inclui etnografia, levantamentos, *record-based analysis* e *history*. O segundo nível, sobre como se analisa os dados, pode ser dividido entre interpretação direta, análise quantitativa e modelagem formal. O terceiro nível, por fim, sobre a quantidade de casos, pode ser dividida em estudo de caso, estudo de poucos casos (*small-N*) e estudo de grande número de casos (*large-N*) (Abbott, 2004).

Parece-nos que o digital transforma especialmente as possibilidades de coleta de dados. O surgimento do conceito de etnografia virtual/digital ou netnografia (Hine, 2008; Kozinets, 1998) empresta o termo e procedimentos da etnografia para dar conta de práticas de observação e participação em espaços de conversações e comunidades online como fóruns. Os levantamentos, pilares da pesquisa quantitativa, mantem uma relativa autonomia de forma, muitas vezes ainda apresentada como a abordagem mais válida. Porém, o modo de abordar os participantes de pesquisa, hoje, usa amplamente de recursos como painéis online, distribuição de questionários por listas de e-mails ou comunidades online e outras

estratégias online (Hesse-Biber & Leavy, 2008; Poynter, 2010). Os outros dois métodos de coleta citados por Abbott (2004), baseados em registros organizacionais/oficiais ou histórico de documentos e pesquisa, tratam de táticas extrativas baseadas na revisão de dados. Nestas, as potencialidades promovidas pelas características da era digital: com o custo negligenciável de armazenamento de dados e a disseminação de dispositivos móveis conectados (Lemos, 2007), praticamente qualquer pessoa pode criar seus próprios registros – na maioria dos casos públicos e armazenados por alguma empresa de tecnologia que transforma dados em valor financeiro.

Entre estas empresas de tecnologia, hoje estão o Facebook e seus mais de 1 bilhão de usuários, o Twitter com seus 200 milhões de usuários e inúmeros casos de análise de dados sociais em tempo real, e a Google, que detém algumas das maiores plataformas digitais de mídia e informação, como o YouTube e o seu buscador homônimo. Toda esta abundância de dados criados a partir de pessoas conectadas entre si chama a atenção de pesquisadores de todos os tipos, que buscam compreender fenômenos micro ou macro-sociológicos a partir dos dados que podem ser extraídos, classificados e cruzados a partir das chamadas mídias sociais. Esta mediação digital de grande parte das trocas comunicacionais na contemporaneidade é um dos grandes motivadores do fenômeno que pretendemos observar.

O atual trabalho faz parte de um projeto independente de investigação que busca compreender as características e procedimentos destes novos desenhos de pesquisa que se focam sobre os dados sociais digitais extraídos dos sites de redes sociais (SRSs). Pesquisadores debruçados sobre questões específicas dos fenômenos da comunicação digital estão utilizando a extração de dados e manipulação das bases de dados nos sites de redes sociais para resolver diversos tipos de questões pertinentes à compreensão das dinâmicas sociais e comunicacionais. Entender questões relacionadas a temas como sociabilidade, linguagem e ciberativismo, comuns objetos de pesquisa na área (Recuero, 2009; Amaral & Montardo, 2011) envolve em considerável grau acessar mensagens e discursos espontâneos dos indivíduos.

Entre o científico e mercadológico, as relações de troca de *expertises*, de um lado, por acesso aos dados, de outro, resultam em pesquisas realizadas em parcerias explícitas – ou mediadas por dispositivos tecnológicos, uma vez que APIs podem ser acessadas de modo direto - de pesquisadores acadêmicos com as bases de dados destas plataformas comerciais, como resultado indireto – *by-product* – das interações cotidianas. Os sites de redes sociais tornam-se então um grande repositório mediador de uma multiplicidade de interesses e olhares sobre as mensagens em circulação.

O valor dos dados sociais extraídos está em constante negociação e os parâmetros e referenciais de cada campo são bastante diferentes, mas algumas estratégias de coleta de dados extrativa ou recrutamento de participantes se aproximam das táticas mercadológicas. A suposta representatividade – graças à pretensa espontaneidade durante a emissão - dos dados sociais digitais na construção de diversos sub-mercados da comunicação corporativa e marketing, como o do “monitoramento de mídias sociais”, termo utilizado para dar conta da coleta, é um destes exemplos. Esta representatividade, por padrões científicos, deveria ser sempre posta em cheque e problematizada, como alertam diversos autores (Freelon, 2012; Bruno, 2012).

Parte das abordagens científicas sobre os métodos digitais problematizam os procedimentos, valores, técnicas e posicionamentos dos dados e suas tecnologias de extração, comunicação e visibilidade. Fala-se de uma redistribuição da pesquisa social promovida, em grande parte, pela “geração rotineira de dados sobre a vida social enquanto parte da própria vida social” (Marres, 2012). A amplitude da produção de dados sociais, assim como o seu potencial de ubiquidade (através do barateamento de *smartphones*, sensores, câmeras, óculos inteligentes e afins) é confrontada com o papel da academia e dos sociólogos, como Marres (2012) que explora o conceito de redistribuição pra apontar como a pesquisa social se reconfigura devido à tais tecnologias, especialmente no que tange suas

delimitações: quem faz, quem usa e quem gera dados, assim como os modos pelos quais resultados circulam na sociedade, mídia e imprensa.

O maior volume dos dados sociais não está mais nas mesmas mãos: se alguns séculos (ou mesmo décadas) atrás, o grande centralizador de dados sociais era o Estado (ou Igreja), hoje organizações comerciais possuem semelhante ou maior quantidade de dados. A atual pertinência do termo *big data* (Mayer-Schonberger & Cukier, 2013) se dá justamente pelas possibilidades de registro e rastreamento que mídias sociais, sensores, mapas, cartões e outras tecnologias de registro e medição adicionaram.

Provocado pelas questões acima, o presente trabalho é um primeiro esforço para responder uma questão de pesquisa central: como os pesquisadores brasileiros da área da Comunicação tem abordado o uso dos dados interacionais em sites de redes sociais?

Metodologia e Questões de Pesquisa.

Considerando que a maioria dos sites de redes sociais populares nos últimos anos foi lançada a partir de 2004, como Facebook (2004), Orkut (2004) e Twitter (2004), decidimos analisar os últimos 10 anos de artigos publicados no NP Tecnologias da Informação e da Comunicação, no GP Cibercultura e no GP Conteúdos Digitais e Convergências Tecnológicas, disponíveis nos anais do Congresso Brasileiro de Ciências da Comunicação, o Intercom, um dos principais eventos da área no Brasil.

Os anais são mantidos em sites específicos de cada edição, que puderam ser acessados a partir do Portal Intercom². Entre 2005 e 2008, os trabalhos ligados à área de Cibercultura faziam parte do Núcleo de Pesquisa em Tecnologias da Informação e da Comunicação. A partir de reativação ocorrida em 2009, trabalhos associados a esta área passaram a ser

distribuídos em dois Grupos de Pesquisa: Cibercultura; e Conteúdos Digitais e Convergências Tecnológicas.

A extração da listagem dos artigos foi apoiada pela ferramenta de *webscraping* Import.io, permitindo a construção de tabelas para análise das variáveis observadas. No total, durante os 10 anos, foram publicados 704 artigos entre os três núcleos/grupos de trabalho. A primeira etapa de recorte filtrou quais destes artigos se baseiam em dados interacionais e opinativos extraídos de sites de redes sociais e fóruns. Optou-se por este recorte por ser, acreditamos, o tipo de informação abundante nos SRSs mais disruptivo em relação à metodologias anteriores de pesquisa empírica na comunicação.

Deste modo, trabalhos que se basearam apenas na coleta de conteúdo de perfis corporativos, jornalísticos ou de personalidades públicas (como políticos e artistas) não foram levados em consideração, uma vez que são mensagens inscritas em planejamentos da recepção mais rígidos, com públicos imaginados definidos e objetivos específicos.

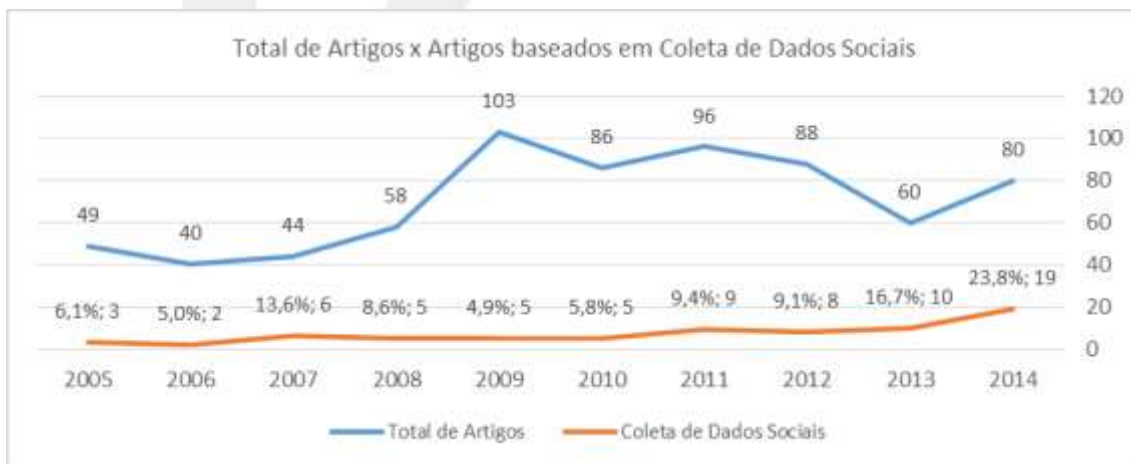
O filtro resultou em um total de 72 artigos, que correspondem a 10,2% da produção do universo selecionado para o estudo. A partir da leitura dos artigos, oito variáveis foram analisadas, permitindo a classificação dos trabalhos com o objetivo de responder às seguintes Questões de Pesquisa:

- a) Quantos trabalhos publicados no evento utilizaram coleta de dados sociais digitais?
- b) Quais sites de redes sociais são abordados nos trabalhos do evento?
- c) Quais as ferramentas utilizadas para coleta e processamento dos dados?
- d) Quais os tipos de conteúdo e volume analisados nos artigos?
- e) Como os pesquisadores abordam a relação entre conteúdo e indivíduos publicadores?
- f) O conteúdo é apresentado de forma direta ou através de visualizações?
- g) Os indivíduos observados, direta ou indiretamente, nos dados empíricos, são anonimizados?

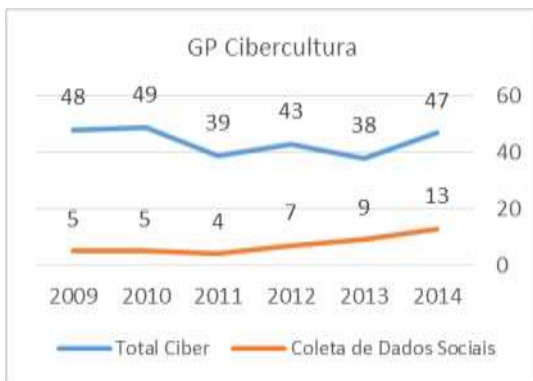
Resultados

Artigos Baseados em Dados Sociais Digitais

O crescimento da relevância dos sites de redes sociais nos últimos anos impactou as modalidades de pesquisa, como pode ser visto na fatia dos artigos com coleta extrativa de dados em sites de redes sociais.



Como podemos ver no Gráfico 1, o número de artigos baseados em dados sociais empíricos digitais cresceu ao longo dos anos de forma irregular, mas chegando a representar quase 1/4 da produção publicada nos anais de 2014 do Congresso. A queda entre os anos de 2008 e 2009 pode ser explicada pela reativação, que trouxe um aumento relevante no número de artigos publicados nos GPs selecionados, especialmente o de Conteúdos Digitais e Convergências Tecnológicas, onde estudos de jornalismo digital e TV digital foram frequentes. No GP de Cibercultura, por sua vez, o crescimento foi acentuado entre 2009 e 2014.

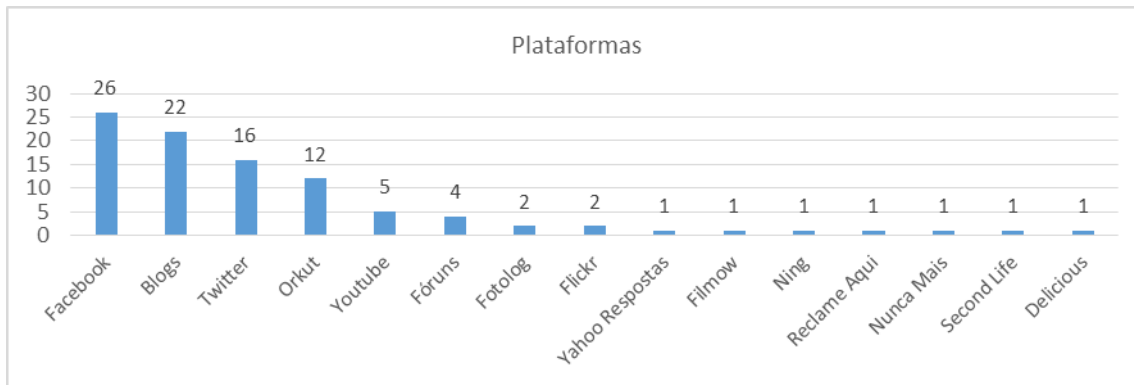


Para fins de comparação, decompomos as taxas do GP de Cibercultura e do GP de Conteúdos Digitais e Convergências Tecnológicas nos gráficos acima.

Pudemos observar que os trabalhos exploram uma amplitude de possibilidades, como foco sobretudo em métodos relacionados a etnografia (referenciada como etnografia virtual ou netnografia), análise de conteúdo e discurso. Uma observação relevante foi a presença de artigos utilizando métodos mistos, especialmente o levantamento (survey) como ferramenta de checagem. Os dados empíricos espontâneos são cruzados com as declarações de indivíduos pré ou posteriormente observados digitalmente, especialmente blogueiros.

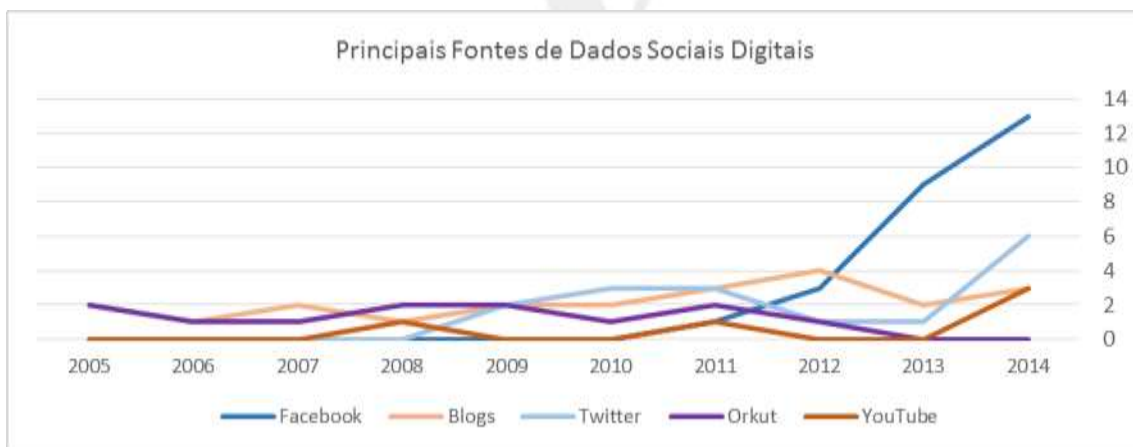
Plataformas Pesquisadas.

As *affordances* dos sites de redes sociais quanto à interação e circulação dos dados somam-se à sua popularidade para a definição de objetos de pesquisa. Ao longo dos 10 anos analisados, Facebook, Blogs, Twitter, Orkut e YouTube lideraram em número de artigos:



O Facebook, site de redes sociais líder no mundo, com mais de 1 bilhão de usuários, chegou à frente em número de trabalhos, com 26 artigos utilizando dados e conteúdo extraídos das interações cotidianas dos usuários através de comentários e postagens. Em segundo lugar, Blogs com 22 trabalhos analisando tanto as plataformas quanto o conteúdo e conversação circuladas. Em terceiro a plataforma Twitter, com 16 artigos analisando perfis e *tweets* neste site de rede social.

Adicionando a variável temporal ao volume de artigos publicados com dados dos cinco principais sites de redes sociais, podemos observar flutuações ao longo dos anos:



O Facebook deu um salto nos últimos anos, para ganhar a liderança isolada como fonte de dados. O Orkut, que reinava como principal site de rede social no Brasil, não foi fonte de nenhum artigo no recorte durante 2013, o que se repetiu em 2014, ano de sua descontinuação pelo Google. Blogs e Twitter flutuam, mas mantem determinada

regularidade. O YouTube, por sua vez, foi fonte de 3 artigos em 2014, enquanto, anteriormente, fora fonte somente de 2 artigos (em 2008 e 2011).

Unidades de Conteúdo e Análise.

A complexidade das plataformas de comunicação digital traz também dúvidas sobre quais unidades de conteúdo são observadas pelos pesquisadores com fins de gerar conhecimento. Entre os SRSs observados, há diferentes tipos, formatos, mídias e até terminologias sobre o que é percebido como um “perfil pessoal”, quais são as unidades de conteúdo trocadas (mensagens, posts, comentários, tweets etc).



Observando as unidades de conteúdo analisadas pelos pesquisadores é possível esboçar uma categorização:

- ✓ **Discursivo-Interacional:** tipo mais comum, que engloba Postagens, Comentários, Tweets e Retweets. É a extração e análise de unidades de conteúdo geralmente textuais de caráter opinativo e conversacional;
- ✓ **Expressivo-Descritivo:** unidades de conteúdo que se referem a processos e recursos de expressão de identidade dos indivíduos, tais como Auto-Descrição nos Perfis e Blogs;
- ✓ **Interacional-Reativo:** trata-se de unidades de conteúdo relativas a opções de feedback do usuário à conteúdos através de recursos simples da interface, sem

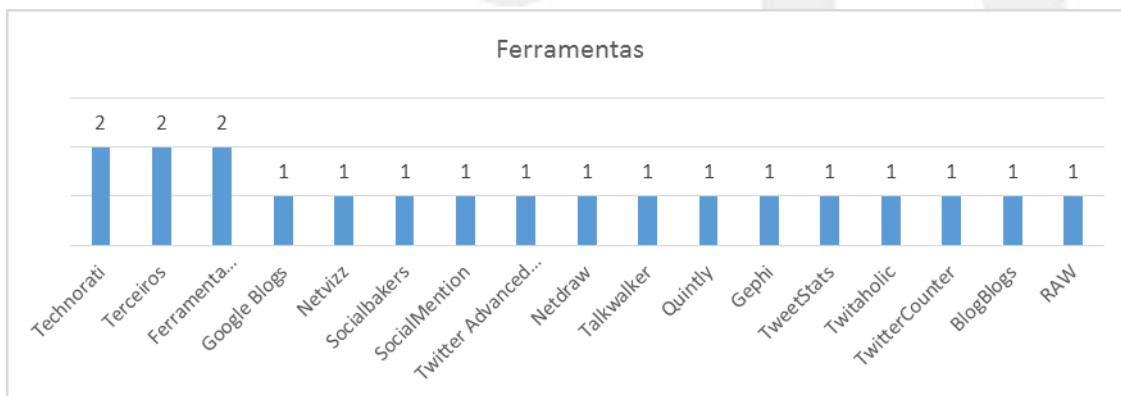
necessidade de elaboração de discurso, tais como uso dos botões Like/Curtir e Share/Compartilhar do Facebook.

Ferramentas e Volume de Coleta.

Talvez uma das questões de pesquisa mais interessantes para os propósitos do projeto do qual o presente artigo faz parte trata das ferramentas utilizadas e volume de coleta de dados. Do total de 72 artigos analisados, apenas 14 (cerca de 19%) utilizaram ferramentas para coleta, armazenamento, processamento ou visualização dos dados. Os 58 restantes foram realizados de forma Manual ou não explicitada.

Em alguns casos certos tipos de ferramentas de exploração mais gerais (como buscadores tais como Google e Bing) parecem estar implícitos nos procedimentos descritos. O mesmo acontece com softwares de planilhas estatísticos. Mas, além de não serem considerados dignos de nota pelos próprios pesquisadores, nosso interesse está especificamente em ferramentas próprias para os métodos digitais de pesquisa.

No gráfico abaixo podemos ver as ferramentas citadas nos 14 artigos que utilizaram e mencionaram explicitamente ferramentas de coleta, processamento ou visualização dos dados. Oito dos artigos utilizaram apenas 1 ferramenta, dois artigos apresentam 2 ferramentas e outros dois artigos citam 4 ferramentas cada.

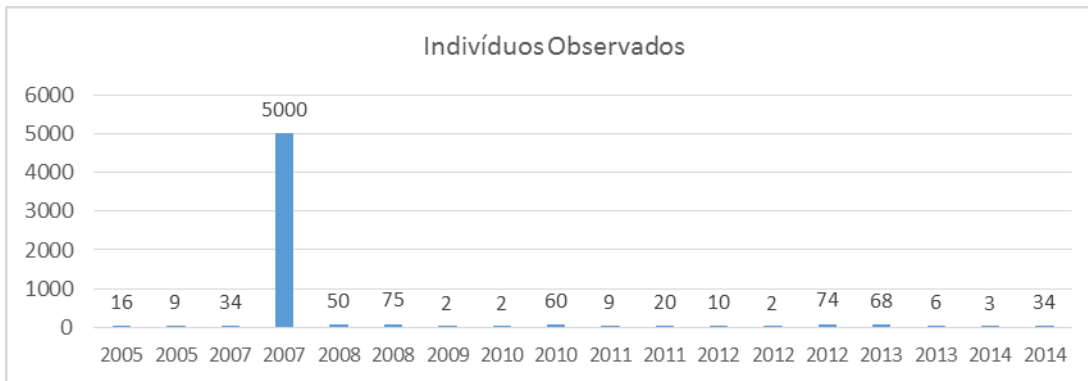


Podemos dividir as ferramentas em seis tipos principais:

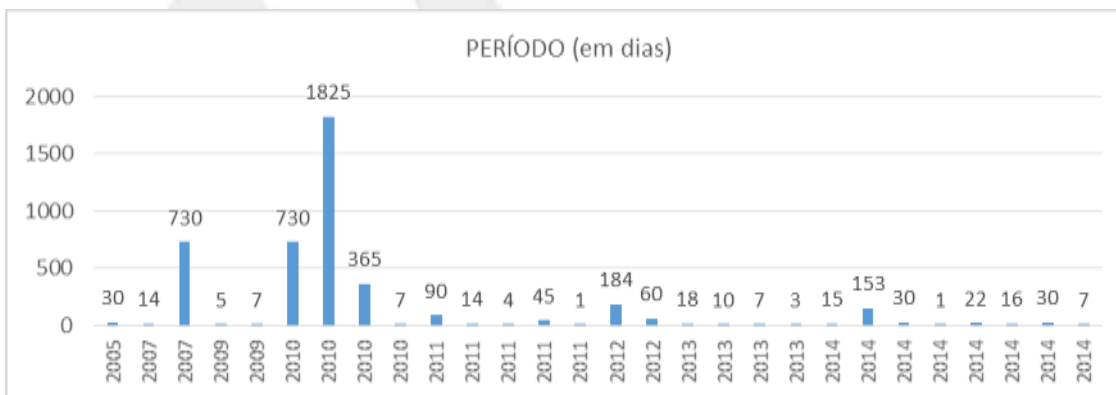
1. Indexadores/Buscadores focados em blogs e mídias sociais, como o Technorati, Blog Blogs, Twitter Advanced Search e o Google Blogs ;
2. Ferramentas comerciais de métricas e performance de mídias sociais: Socialbakers, Quintly, TweetStats, Twitaholic e TwitterCounter;
3. Ferramentas comerciais de Monitoramento de Mídias Sociais: Talkwalker e SocialMention;
4. Ferramentas próprias: crawlers e softwares desenvolvidos para a pesquisa, como o Neofluxo;
5. Ferramentas de coleta e visualização de redes: Gephi, Netvizz e Netdraw;
6. Ferramentas de visualização de dados: Raw.|

Parece predominar, nos trabalhos analisados, uma postura de observação dos fenômenos socio-comunicacionais nos SRSs como objetos a serem acessados e confrontados de modo direto, sem transformação ou processamento dos itens em indicadores ou pontos de dados categóricos ou sistematizados. Apesar da frequente referência a métodos etnográficos digitais, predomina a baixa média de dias de pesquisa/observação e pequeno envolvimento com os participantes, pilares do método.

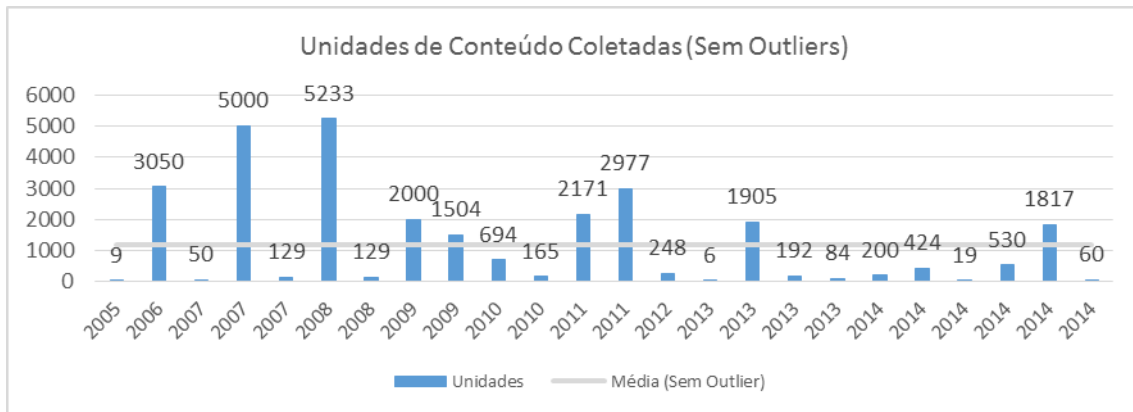
Em apenas 17 dos trabalhos foi declarada a quantidade de indivíduos (através de perfis, blogs etc) observados. Variou de 2 a 5000, resultando numa média de 321 pessoas. Excluindo-se da conta um *outlier* isolado de 5000 indivíduos, a média cai para 29 pessoas.



Entre os 28 artigos que trouxeram a informação do período de coleta e/ou observação dos SRSs, a média de dias foi de 157. Vale observar também que o número médio no ano chegou a 372 em 2007 e 731 em 2011, mas foi de apenas 35 dias em 2014.



Em relação ao volume, 26 dos trabalhos declararam a quantidade de conteúdo (mensagens, comentários, posts etc) coletada e/ou analisada. O gráfico abaixo apresenta o volume com a exceção de dois *outliers*: artigo que utilizou dados a partir da coleta de 340 milhões de posts no Facebook; outro artigo que se baseia na exploração de 20 milhões de mensagens em diversos SRSs em torno da temática eleitoral.

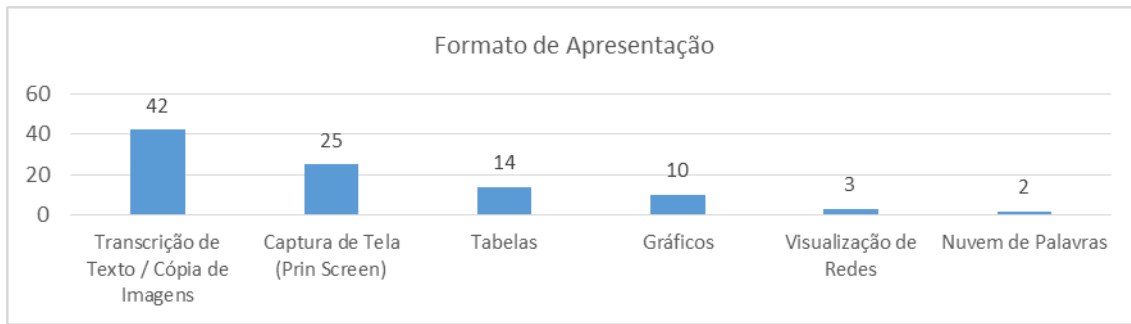


Com exceção dos dois outliers citados anteriormente, o volume variou de 6 a 5233 unidades de conteúdo coletadas/analizadas. A linha horizontal no gráfico acima representa a média (excluído os dois valores de cada extremidade da série no cálculo): 1299,1. Este número é especialmente interessante ao percebermos que boa parte das ferramentas comerciais de agregação de dados sociais digitais oferecem trials que normalmente superam este volume de coleta. Em trabalho no qual David Beer apresenta uma interessante exploração de um destes *softwares*, aponta também como a própria relação do pesquisador com estes softwares é digna de observação:

the a priori analytics occurring in this software need to be fully appreciated where possible. [...] We are faced with the presence of the power of algorithms in our method, these algorithms will come to shape the outcomes of social research methods as they sort-out our data and present it to us in a manageable form (Beer, 2012, p. 10).

Visualização.

Outra questão de pesquisa que instigou o levantamento se refere aos tipos de visualização utilizadas para apresentar os resultados. Identificamos sete tipos de exibição dos dados: Transcrição e Cópia de Imagens; Capturas de Tela (Print Screen); Tabelas; Gráficos; Visualização de Redes; e nuvem de Palavras.

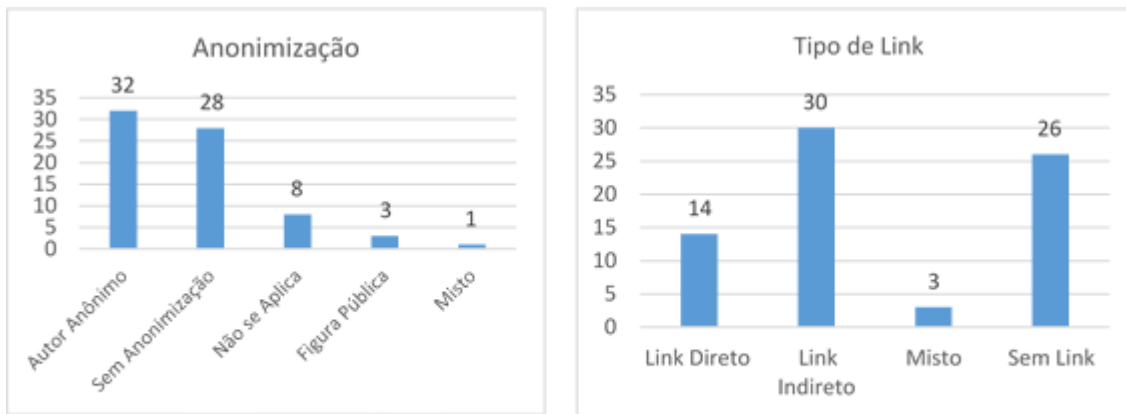


A Categoria que liderou foi a de “Transcrição e Cópia de Imagens”, que significa a cópia do conteúdo em formato textual ou a reprodução de conteúdos do tipo imagem (como *image macros* ou *memes*) analisados. A Captura de Tela representa um modo diferente de apresentação do conteúdo e mensagens, por trazer também o contexto de interface da plataforma, como tipografia, botões, layout etc. As Tabelas e Gráficos, por sua vez, foram utilizados especialmente quando existe a adição de camadas de informação, como variáveis temporais ou codificação de categorias de análise.

Particularizamos os itens “Visualização de Redes” e “Nuvem de Palavras” por motivos semelhantes. A visualização de redes é uma técnica metodológica já com cerca de 100 anos de aplicação nas ciências sociais, mas ganhou novo fôlego com a estrutura de rede de plataformas online como as pesquisadas. O mapeamento das conexões entre os indivíduos e suas presenças online agrega dados relacionais, expandindo a compreensão dos dados de atributos e análise de conteúdo.

As nuvens de palavras, por sua vez são representações gráfico-visuais da frequência de ocorrências de termos em um texto, visualização na qual o próprio tamanho relativo de cada palavra representa sua frequência. Como apontou Manovich (2011, p.158), o formato é um tipo de visualização direta, no qual “os dados são reorganizados em uma nova representação visual que preserva sua forma original”. Enquanto um dos formatos que aproveita as vantagens computacionais e formato digital ao prescindir do princípio básico de redução nos tipos clássicos de visualização, a baixa frequência de utilização é um dado digno de nota.

Anonimização e Referenciação.



Por fim, as últimas variáveis analisadas tratam da presença ou ausência de anonimização dos conteúdos apresentados, assim como o nível de referenciação destes conteúdos, item estritamente relacionado à capacidade do leitor em descobrir indiretamente os emissores. Cerca de 44% dos artigos utilizaram algum tipo de Anonimização contra 39% que não tiveram o mesmo cuidado.

Quanto à referenciação de plataformas e URLs, 14 trabalhos ofereceram links diretos aos conteúdos analisados. Em parte dos casos, pode ser considerada uma prática potencialmente nociva de exposição dos indivíduos observados. Em 30 trabalhos foram publicados Links Indiretos, quando o link vai por exemplo, para a comunidade observada mas não para o tópico específico. Em três trabalhos houve a ocorrência das duas modalidades e em 26 não foram oferecidos links.

Conclusões e Próximos Passos.

O presente trabalho conseguiu demonstrar a relevância do número de trabalhos da área de Cibercultura que se baseiam em dados empíricos coletados de sites de redes sociais. Nos 10 anos de produção publicada nos anais dos GPs escolhidos, ocorreu um crescimento da metodologia de investigação, indicando ser uma temática que merece a atenção de

interessados nos processos de produção científica do conhecimento. O aumento da oferta de dados interacionais e opinativos dos usuários leva interessados a temáticas como opinião pública, consumo, marketing boca-a-boca e sociabilidade a abraçar ambientes digitais como essencial arena de entendimento tanto da própria internet como da sociedade como um todo. Parece que “a questão não é mais ver o quanto da sociedade e cultura está online, mas também como diagnosticar mudanças culturais e condições da sociedades através da Internet” (Rogers, 2013, p.21).

Resgatar e problematizar, além de contrapor às novas possibilidades e desafios dos sites de redes sociais, os procedimentos de coleta, análise e apresentação dos resultados de pesquisa nestes ambientes demonstra-se como um objetivo relevante e com potenciais frutos para os perfis de interesses de pesquisa que foram descobertos no estudo. Nas ferramentas apresentadas neste trabalho e nas que se fazem ausentes aqui, sobreposições entre metodologias com referências científicas ou vinculadas à mercado complexificam o ecossistema dos softwares capazes de apoiar os pesquisadores. Como próximos passos, o presente planejamento de pesquisa procurará se debruçar em outros campos acadêmicos e mercadológicos de exploração dos dados sociais digitais. Os diagnósticos levantados no trabalho apontam para a validade dos objetivos de problematização dos sites de redes sociais, ambientes ricos de informação mas comerciais, quanto suas práticas científicas e comerciais de uso destes dados, visando a produção de mapeamentos e materiais reflexivos-instrucionais sobre as práticas de exploração dos SRSs.

Bibliografía.

Abbott, Andrew. (2004). *Methods of Discovery: Heuristic for the Social Sciences*. New York (EUA): W. W. Norton & Company.

Amaral, A. R. ; Montardo, S.P. (2011). Pesquisa em Cibercultura: análise da produção científica brasileira na Intercom. In: *Revista Logos*, UERJ, Rio de Janeiro, Ed. 34, v. 18, p. 8.

Beer, David. (2012). 'Using social media data aggregators to do social research'. In: *Sociological Research Online*, v. 17, n.3.

Bruno, Fernanda. (2012). *Rastros Digitais: o que eles se tornam quando vistos sob a perspectiva da teoria ator-rede?*. *Anais do Encontro da Compós*.

Freelon, Deen. (2013) On the interpretation of digital trace data in communication and social computing research. In: *Journal of Broadcasting & Electronic Media*.

Hesse-Biber, Sharlene; Leavy, Patricia. (2008). *Handbook of Emergent Methods*. Londres: The Guilford Press.

Hine, Christine. (2008). "Virtual Ethnography." In GIVEN, Lisa. *The SAGE Encyclopedia of Qualitative Research Methods*, 922-25, Thousand Oaks, CA: SAGE Publications, Inc.

Kozinets, Robert. (1998) On Netnography: initial reflections on consumer research investigations of cyberculture. *Advances in consumer research*, v. 25, n.1, pp.: 366-371.

Lemos, André. (2007). Comunicação e práticas sociais no espaço urbano: as características dos Dispositivos Híbridos Móveis de Conexão Multiredes (DHMCM). *Comunicação, Mídia e Consumo* (São Paulo), v. 4, p. 23-40.

Manovich, Lev. (2011). O que é visualização? Trad. Ana Elisa Ribeiro, Francis Arthuso Paiva e Vinícius Martins Rocha. *Estudos em Jornalismo e Mídia*, v. 8, n. 1, p. 146- 171..

Manovich, Lev. (2012). Trending: the promises and the challenges of big social data. In: GOLD, Matthew K. (ed.) Minnesota (EUA): *Debates in the Digital Humanities*. The University of Minnesota Press.

Mayer-Schonberger, Viktor. (2009). *Delete: the virtue of forgetting in the digital age*. Princeton, Princeton University Press.

Marres, Noortje. (2012). The redistribution of methods: on intervention in digital social research, broadly conceived. In: *Sociological Review*, vol. 60, s. 1, 2012. pp. 139-165.

Poynter, Ray. (2010). *Handbook of Online and Social Media Research*. West Sussex (Reino Unido): Wiley.

Recuero, Raquel. (2009). *Redes Sociais na Internet*. Porto Alegre: Sulina, 2009.

Rogers, Richard. (2013). *Digital Methods*. Londres: The MIT Press.

Savage, Mike; Burrows, Mike. (2009). Some Further Reflections on the Coming Crisis of Empirical Sociology. *Sociology*, vol. 43, n. 4, pp. 765–775.

Scorali, Carlos Alberto. (2009). Mapping conversations about new media: the theoretical field of digital communication. *New Media Society*, 11, pp.943-964.

¹ Tarcízio Silva é mestre em Comunicação e Cultura Contemporâneas pela Universidade Federal da Bahia. Currículo Lattes: <http://lattes.cnpq.br/8992630383672112>; Blog: tarcizosilva.com.br/blog; Email: eu@tarcizosilva.com.br

² http://www.portalintercom.org.br/index.php?option=com_content&view=article&id=1081&Itemid=125