

ACESSO AOS DADOS DO FACEBOOK E O JORNALISMO INDEPENDENTE NA AMÉRICA LATINA

[ARTIGO]

Daniela Osvald Ramos

Universidade de São Paulo. Escola de Comunicações e Artes

Egle Müller Spinelli

Escola Superior de Propaganda e Marketing

Mário Arruda

*Universidade Federal do Rio Grande do Sul.
Faculdade de Biblioteconomia e Comunicação*

[RESUMO ABSTRACT RESUMEN]

Compreender o acesso aos dados das mídias sociais é de extrema relevância, principalmente quando um dos maiores serviços de mídia social, o Facebook, altera constantemente suas APIs (Application Programming Interface), a Interface de Programação de Aplicativos que permite a extração de dados da plataforma. Este estudo mostra os desafios e dificuldades de ter acesso aos dados para explorar o alcance das conexões e associações nas páginas de iniciativas de jornalismo digital na América Latina. Para esta verificação, pretendeu-se realizar a mineração de dados no Facebook em três veículos independentes latino-americanos – Jota (Brasil), Meio (Brasil) e Nómada (Guatemala) – a partir de três ferramentas: o Netvizz, aplicativo locado no Facebook para extração de dados; o software livre Gephi para gerar gráficos e análises de rede; e o aplicativo pago Social Monitor, que monitora o engajamento nas redes sociais.

Palavras-chave: Facebook. Dados. Jornalismo Independente. Transparência. América Latina.

The access to social media data has become extremely relevant, especially when one of the largest social media services, Facebook, constantly changes its APIs (Application Programming Interface), that allows the extraction of data from the platform. This study shows the challenges and difficulties of accessing data to explore the reach of connections and associations on the pages of digital journalism initiatives in Latin America. For this verification, it was intended to perform data mining on Facebook in three independent Latin American vehicles – Jota (Brazil), Meio (Brazil) and Nómada (Guatemala) – from three tools: Netvizz, a data collection and extraction application on Facebook; Gephi, a open source software to generate graphs and network analysis, and the paid application called Social Monitor, which monitors social networking engagement.

Keywords: Facebook. Data. Independent Journalism. Transparency. Latin America.

La comprensión del acceso a los datos de los medios sociales es de extrema relevancia, principalmente cuando uno de los mayores servicios de medios sociales, Facebook, cambia constantemente sus API (Application Programming Interface), la Interfaz de Programación de Aplicaciones que permite la extracción de datos de la plataforma. Este estudio muestra los desafíos y dificultades para acceder a los datos para explorar el alcance de las conexiones y asociaciones en las páginas de iniciativas de periodismo digital en América Latina. Para esta verificación, se pretendió realizar la minería de datos en Facebook en tres vehículos independientes latinoamericanos – Jota (Brasil), Medio (Brasil) y Nómada (Guatemala) – a partir de tres herramientas: el Netvizz, aplicación de Facebook para la extracción de datos; el software libre Gephi para generar gráficos y análisis de red; y la aplicación pagada Social Monitor, que monitorea el compromiso en las redes sociales.

Palabras clave: Facebook. Datos. Periodismo Independiente. Transparência. América Latina.

Introdução

Iniciamos a pesquisa para este artigo com a seguinte indagação: “É possível entender a atuação das organizações jornalísticas no Facebook a partir dos dados gerados pelas suas páginas?” Para tanto, foram eleitas três marcas de jornalismo independente com base no Anuário Brio de Jornalismo 2018. Dentre os cinco projetos em destaque, foram selecionados os brasileiros Jota¹ e Meio² e o guatemalteco Nómada³, por estabelecerem estratégias digitais em diferentes plataformas, terem presença no Facebook e serem representantes latino-americanos. O Jota existe desde 2014 e é um site que oferece informações exclusivas, análises e notícias de bastidores jurídicos. O Meio surgiu em 2017 e é uma newsletter distribuída por e-mail de segunda a sexta que traz as principais notícias do dia para serem lidas rapidamente no ambiente digital. O Nómada foi lançado em 2014, na Cidade da Guatemala, e é um meio jornalístico que aborda política e cotidiano.

A partir dessas três iniciativas, pesquisamos qual seria a metodologia de análise de redes sociais possível de ser aplicada nesses casos. Nessa etapa inicial, por meio do uso de ferramentas gratuitas

que extraem dados das páginas do Facebook e geram grafos – Netvizz e Gephi –, foi constatado que tanto o Jota como o Meio não davam total acesso aos dados que pudessem indicar percepções de padrões de uma rede por meio de estruturas de visualização de dados. Apenas o projeto Nómada estava aberto⁴ para a mineração de dados na rede social Facebook. Com isso, surgiu um questionamento para a pesquisa – da proposta inicial, que era comparar e analisar dados a partir da formação de rede (grafos), partiu-se para a indagação sobre como analisar dados extraídos com o auxílio de diferentes ferramentas, já que o acesso ao *Application Programming Interface* (API) do Facebook⁵ está cada vez mais restrito.

A questão do monopólio do Facebook sobre as informações na rede e o recente escândalo de venda de dados de domínio privado extraídos do Facebook pela Cambridge Analytica culminaram na diminuição do acesso, antes público, de níveis de dados. O acesso aos dados foi se restringindo aos poucos na história do Facebook. Por exemplo, em 30 de abril de 2015, a rede social já tinha descontinuado algumas funcionalidades de sua *Graph Application Programming Interface* ao extinguir a função de procura por palavras-chave (*search*), que retornava postagens e comentários de perfis e de páginas que correspondiam a termos determinados (ALVES, 2016, p. 68). Em janeiro de 2018, uma mudança de algoritmo passou a privilegiar mensagens de familiares e amigos

[1] JOTA: notícias que fazem diferença, c.2014. Informações institucionais de qualidade: Justiça, STF, STJ, TJs, Carf, Cade, agências reguladoras e Legislativo. Notícias, bastidores e análise. Disponível em: <<https://www.jota.info/>>. Acesso em: 2 Ago. 2018.

[2] CANAL MEIO S.A. **Meio**, c2016-2018. Apresenta notícias do Brasil e do mundo. Disponível em: <<https://www.canalmeio.com.br/>>. Acesso em: 8 jan. 2018.

[3] NÓMADA, c2014. Apresenta notícias sobre política e cotidiano. Disponível em: <<https://nomada.gt/>>. Acesso em: 22 fev. 2018.

[4] Extração de dados feita em 29 de março de 2018.

[5] GRAPH API. In: Facebook, c2018. Disponível em: <<https://developers.facebook.com/docs/graph-api/>>. Acesso em: 29 mar. 2018.

na *timeline* do Facebook, fazendo com que o conteúdo informativo ficasse em segundo plano na mediação do algoritmo, a não ser que fossem comentadas por pessoas da rede do usuário. Desde o dia 6 de fevereiro de 2018, o Facebook parou de disponibilizar informações dos usuários que realizam comentários nas publicações das páginas. Ou seja, a partir de agora não é mais possível extrair informações como nome, foto e ID do usuário, ao menos que você tenha um *token* de CRM⁶ da página (AVISO, 2018).

Em 4 de abril de 2018 o Facebook, por causa do vazamento de dados de usuários relacionado ao caso da Cambridge Analytica, passa a fazer constantes alterações nas APIs⁷ de acesso aos dados da plataforma para garantir a segurança e privacidade dos usuários como política corporativa, o que implica na descontinuação de algumas funcionalidades de plataformas e ferramentas de monitoramento de dados nessa rede social.

Nesse contexto, questões que envolvem opacidade, transparência e a natureza dos dados extraídos das mídias sociais passa a ser um assunto importante a ser discutido quando se pensa na democratização das informações para o fortalecimento de processos democráticos.

[6] O CRM (Customer Relationship Management) é um gerenciador das ferramentas de comunicação com o cliente. Apenas com a permissão do administrador da página do Facebook é possível ter acesso aos dados dos usuários.

[7] É importante compreender que a API é, por um lado, um conjunto de comandos técnicos para realizar ações, mas de outro é uma documentação que representa as políticas empresariais do Facebook para acesso ao seu banco de dados (ALVES, 2016, p. 75).

Pela dificuldade de aplicação das ferramentas gratuitas (Netvizz e Gephi) inicialmente selecionadas para a pesquisa de análise dos casos selecionados, optou-se em apresentar neste artigo uma discussão das possibilidades de análise dos dados abertos dessas marcas no decorrer do estudo, a partir de ferramentas que eram acessíveis aos pesquisadores na época das análises, mostrando um pouco da diversidade metodológica possível para a análise de redes sociais. Verificou-se que, com as últimas mudanças na API do Facebook, as ferramentas selecionadas para esta pesquisa (Netvizz e Gephi) não conseguiam mais ter acesso aos dados das páginas selecionadas, apenas aos da página Nómada, que teve a extração de dados realizada antes da mudança. Dessa maneira, optou-se em utilizar uma ferramenta paga e acessível aos pesquisadores, denominada Social Monitor⁸, para análise de audiência e verificação de estratégias digitais das iniciativas jornalísticas Jota e Meio.

1. Opacidade e transparência de dados

O contexto da opacidade e transparência de dados gerados pelos usuários de formatos de mídias sociais como o Facebook vai além da discussão pretendida aqui e se expande para uma reflexão sobre o uso de *big data* para a criação de modelos de

[8] SOCIAL MONITOR. Disponível em: <<http://www.socialmonitor.com.br/>>. Acesso em: 23 mar. 2018.

sistemas diversos que podem ser usados por governos, empresas e diversas instituições, às vezes de forma a aumentar a desigualdade e ameaçar a democracia, como aponta O'Neil (2016). A autora, uma matemática que trabalhou no mercado financeiro e o abandonou após a crise de 2008, dedica-se a explicar de que elementos e como é programada a caixa preta de análises, com funções geralmente preditivas e de padrões de *big datas* modelados por algoritmos diversos.

Um elemento crucial para o tratamento de dados é a definição de um modelo, definido por ela como uma representação abstrata de um processo (O'NEIL, 2016, p. 18), através do qual o processo de análise se dará. O'Neil (2016, p. 15-31) descreve três etapas: análise de estatísticas do esporte beisebol; modelo pessoal; e o modelo do *recidivism risk* ("risco de reincidência"), este último aplicado no sistema penal estadunidense. O primeiro é baseado em estatísticas públicas rigorosas das temporadas esportivas, o que deixa o modelo dinâmico e com a possibilidade de ser acompanhado publicamente. É um exemplo de um modelo que utiliza um *proxy*⁹ público no tratamento dos dados, ou seja, os dados e as possíveis correlações efetuadas entre eles são de domínio público, transparentes.

Mas seu foco é o reconhecimento do que chama *weapons of math destruc-*

tion ("armas de destruição matemática", em tradução livre), ou seja, modelagens de análise de dados que necessariamente contêm os seguintes elementos: opacidade no tratamento dos dados, escala de aplicabilidade e dados que causam algum tipo de dano. Como um dos exemplos, ela analisa um modelo utilizado pelo sistema de segurança penal nos Estados Unidos, o *recidivism risk*, que mede um provável índice de reincidência criminal (O'NEIL, 2016, p. 23-7), e que já afeta a população negra e hispânica daquele país. Entre seus componentes, está um questionário que os detidos à espera de julgamento precisam responder na prisão. No entanto, os próprios presos não sabem que esse questionário será insumo para o cálculo desse índice de reincidência e também não são informados do peso matemático que cada resposta tem na análise final. Assim, perguntas como "Tem familiares com antecedentes criminais?" afeta muito mais a família de negros e hispânicos do que de brancos e produz um alto índice de pessoas dessas origens encarceradas por mais tempo.

O outro modelo descrito seria um construído de acordo com as necessidades pessoais para resolver um problema específico, por exemplo, "o que cozinhar para a minha família quando eu estiver ausente?". O'Neil (2016, p. 20) descreve como construiria esse modelo, que critérios usaria, em que medida estes seriam respeitados, ou seja, as variáveis, e deixa claro que, de certa forma, um modelo sempre vai refletir, principalmente em seus pontos cegos, os julgamentos e prioridades do seu criador. Mas e quando seu criador é uma empresa gigante da web, como o Google, Amazon ou Facebook?

[9] Em tradução livre do inglês, *proxy* significa "representante", "autoridade", "procurador". O termo é utilizado na área da informática para designar o elemento que intermedeia as requisições de acesso a conteúdo de um computador pessoal para um outro computador conectado à internet, geralmente de acesso público.

And yet many companies go out of their way to hide the results of their models or even their existence. One common justification is that the algorithm constitutes “a secret sauce” crucial to their business. It’s *intellectual property*, and it must be defended, if need be, with legions of lawyers and lobbyists. In the case of web giants like Google, Amazon, and Facebook, these precisely tailored algorithms alone are worth of billions of dollars¹⁰ (O’NEIL, 2016. p. 29).

Ou seja, para responder a uma das perguntas que nos fizemos quando no início desta pesquisa – “Qual é a natureza da pesquisa em comunicação digital: que dados são estes?” –, chegamos à conclusão parcial de que são dados oriundos de um modelo de opacidade, o que ficou ainda mais evidente no já citado caso do Facebook com a Cambridge Analytics. O’Neil (2016) chama esse modelo de WMD (sigla para *Weapons of Math Destruction*), mas na área do marketing digital é uma estratégia de negócio chamada *walled garden*, ou “jardim murado”: o que acontece naquele território só é de conhecimento do detentor dos muros. *A Walled Garden is a closed ecosystem in which all the operations are controlled by the ecosystem operator* (POULPIQUET, 2017). ¹¹

[10] “E, no entanto, muitas empresas se esforçam para esconder os resultados de seus modelos ou até mesmo sua existência. Uma justificativa comum é que o algoritmo constitui “um molho secreto” crucial para seus negócios. É propriedade intelectual, e deve ser defendido, se necessário, com legiões de advogados e lobistas. No caso de gigantes da web como Google, Amazon e Facebook, esses algoritmos precisamente adaptados valem bilhões de dólares.” Tradução livre dos autores.

[11] “Um Jardim Murado é um ecossistema fechado no qual todas as operações são controladas pelo operador do ecossistema.”, em tradução livre dos autores.

Dessa forma, os dados disponíveis para a pesquisa da distribuição de conteúdos pelos novos modelos de negócio que também são conglomerados de mídia, como a rede social objeto do nosso enfoque, serão sempre parciais. As marcas de jornalismo que analisamos estão dentro do que pode ser chamado de um posicionamento “independente”, mas, para além do conteúdo, o que está em jogo é a democratização e regulação de acesso aos dados gerados pela mídia da qual se usam para propagar seu conteúdo. Nesse ambiente, o modelo algorítmico de tratamento dos dados (*inputs*) dos seus usuários não é levado a escrutínio público. Dessa forma, é o resultado dos *inputs* e das dinâmicas que o Facebook promove entre e a partir deles, como a venda de dados pessoais sem conhecimento prévio do público, cujos resultados nos aparecem como novos fenômenos midiáticos. O exemplo mais recente é a discussão do uso do Facebook nas últimas eleições políticas estadunidenses. Esse resultado é um pouco diferente do que conhecíamos do efeito pelo jogo de forças de influência das publicações tradicionais da imprensa profissional, seja da indústria da comunicação ou independente, pois envolve estratégias de publicação de conteúdo altamente personalizadas, que envolvem análises de um grande volume de dados (*big data*) em tempo real. Esse tipo de estratégia de uso de dados contou com um fator de opacidade para que acontecesse. E o que é transparente para os pesquisadores do campo? É a nossa discussão a seguir, na qual discorreremos sobre as metodologias em curso adotadas para análise das redes sociais digitais e como as aplicamos em nossos objetos, a partir do acesso (ou não) aos dados públicos das marcas disponibilizadas pelo Facebook.

2. Metodologia: análise de mídias sociais

A análise de redes sociais na internet consiste em um trabalho interdisciplinar, uma compilação de métodos estruturados que, segundo as premissas da Análise de Redes Sociais (ARS), parte do princípio de que ao estudar as “estruturas decorrentes das ações e interações entre os atores sociais, é possível compreender elementos a respeito desses grupos e, igualmente, generalizações a seu respeito” (FRAGOSO; RECUERO; AMARAL, 2011, p. 115). Ao iniciar a pesquisa para este artigo, partimos das indagações iniciais – “Qual é o alcance de marcas jornalísticas independentes no Facebook? Qual é a rede, com quem se relacionam?” –, e iniciou-se a escolha de casos que se adequassem a este escopo. Selecionamos três veículos independentes na América Latina – Jota (Brasil), Meio (Brasil) e Nómada (Guatemala) – por terem sido indicadas pelo Anuário Brio de Jornalismo 2018 dentre os cinco projetos que valem a pena acompanhar, principalmente por estabelecerem estratégias digitais que utilizam a rede social como forma de engajar a audiência. Queríamos partir da compreensão dessas iniciativas nas redes sociais como “nós”, conforme definidos por Fragoso, Recuero e Amaral (2011), representados pelos atores envolvidos e suas representações (nesta pesquisa indicados pelas páginas do Facebook) e as possíveis arestas ou conexões que estabelecem, para discorrer sobre ferramentas e softwares *open source* de análise de redes sociais e indicar metodologias quantitativas e qualitativas desse campo.

Extrair dados e gerar visualizações dessas redes é um desafio para o pesquisador, principalmente porque os algoritmos

das redes sociais operam em um terreno nebuloso e são alterados a todo momento conforme os interesses e necessidades dos proprietários das empresas de tecnologia, no caso deste estudo, o Facebook. Antes das últimas mudanças no API, essa plataforma permitia monitorar *fan pages* e grupos públicos, mas sempre dificultou o acesso a grupos privados ou perfis pessoais, mesmo públicos. O Facebook também possibilita que os usuários da plataforma configurem certa privacidade de acesso aos dados, podendo até fechar informações caso não queiram que os dados sejam analisados por terceiros. Dessa maneira, se o usuário ou grupos pesquisados tiverem dados privados ou exclusivos, as ferramentas de monitoramento não conseguem extrair as informações.

Enquanto as organizações adotam a mídia social para a colaboração e comunicação, há uma necessidade crescente de ferramentas que auxiliem a mapear e a medir redes sociais empresariais. Enquanto nossa sociedade adota a mídia social como um novo fórum para o discurso público, criando uma praça virtual, há uma necessidade crescente de ferramentas e métodos que possam documentar esses espaços (RECUERO; BASTOS; ZAGO, 2015, p. 10-1).

Este estudo mostra, a partir da experiência de uso com ferramentas e softwares *open source*, como o aplicativo Netvizz¹² e o software Gephi¹³ e a ferra-

[12] Disponível em: <<https://apps.facebook.com/107036545989762/>>. Acesso em 2/8/2018.

[13] THE GEPHI CONSORTIUM. Gephi: makes graphs handy, c2008-2017. Disponível em: <<https://gephi.org/>>. Acesso em: 28 mar. 2018.

menta paga Social Monitor, possibilidades de análise a partir da característica da extração dos dados e, principalmente, as barreiras e dificuldades encontradas para desenvolver pesquisas científicas na área. Das três iniciativas selecionadas – Jota e Meio (Brasil) e Nómada (Guatemala) – apenas a última apresentava os dados abertos para a análise por meio de ferramentas gratuitas (Netvizz e Gephi), os quais foram extraídos no dia 29 de março de 2018, antes das ações de restrição de acesso aos dados pelo Facebook, ocorridas em 4 de abril de 2018. As páginas no Facebook do Jota e do Meio já apresentavam o fechamento de dados pela forma como foram configuradas pelos próprios veículos, mesmo antes das últimas condições restritivas da rede social. Essa situação ocasionou a busca de alternativas para a análise dos dados dessas páginas, o que foi possível pelo acesso dos pesquisadores à ferramenta Social Monitor, utilizada pelos maiores portais de notícias do Brasil para identificar o engajamento de conteúdo de páginas nas redes sociais, por meio da mensuração de número de likes, posts, comentários, compartilhamentos, reações e novos seguidores. Estas duas páginas foram monitoradas para este estudo durante 19 dias – de 23 de abril a 11 de maio de 2018.

3. Análise dos Casos

No que toca à extração de dados do Facebook a partir do aplicativo Netvizz, o que

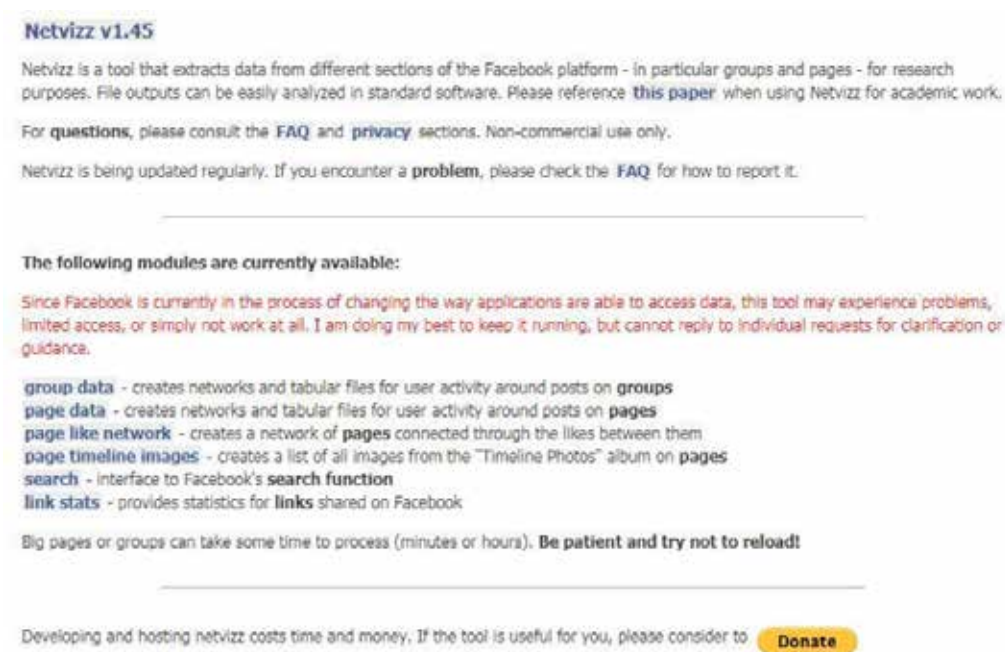
constatamos é que suas possibilidades se encontram extremamente reduzidas. O estágio do próprio aplicativo pode ser visto como um rastro das rápidas transformações resultantes do embate de forças em âmbito comercial, institucional e ético pelo qual passa o Facebook e outros sites de redes sociais. A questão da privacidade de dados, uma preocupação bem mais próxima de uma ética hacker no passado, hoje é um tema central em diferentes estratos políticos.

O efeito de tal embate é o conjunto de restrições pelo qual sofre o Netvizz hoje. Um aplicativo categorizado no Facebook como referente à educação, ele foi largamente utilizado por empresas de publicidade e marketing para visualização de nichos de mercado – era uma chave que abria atalhos para a chegada na mina de ouro de dados personalistas armazenados no Facebook. E de certo modo, isso feria diretamente os próprios termos de uso do site de redes sociais que o abriga: o Facebook promete que não entregará dados pessoais a terceiros, somente servindo como um veículo para mensagens.

Diante disso, a chave-mestra foi ficando cada vez mais restrita, como mostra a Figura 1. Hoje, os nomes dos usuários e sua localidade são mantidos em sigilo. A única ferramenta que ainda é capaz de produzir a visualização de redes é a chamada *page like network*, que se refere à extração de dados que mostram os *likes* que uma página dá em outra. Isso se mantém devido ao caráter público do *like* – interação banal a um primeiro olhar, mas que pode evidenciar interações de concordância ideológica, política, estética, etc.

[Figura 1]

Print screen da interface do aplicativo Netvizz em 27.04.2018



Fonte: Netvizz (2018).

Mas vale salientar também que, seja qual for a extração de dados, não existe garantia de que a imagem gerada seja correspondente à rede presente no Facebook. Isso porque temos sempre uma extração parcial relacionada a apenas um dos quesitos que montam as redes – a frequente mudança dos quesitos de relevância de dados no Facebook não muda apenas a circulação momentânea de dados, mas age sobre a própria rede de páginas e usuários, já que a rede não é estanque, mas se estabelece justamente por suas relações momentâneas orientadas pela circulação.

Isso se comprova ainda mais quando utilizamos o Gephi para montagem de gráficos a partir dos dados extraídos no Netvizz. No Gephi é preciso escolher os quesitos que fazem a forma da imagem, como veremos no tópico seguinte. No entanto, é a partir da visualização dos gráfi-

cos e da análise de como agem os quesitos de relevância em relação à sua forma é que podemos entender as redes formadas a partir do complexo gerenciamento de dados do Facebook.

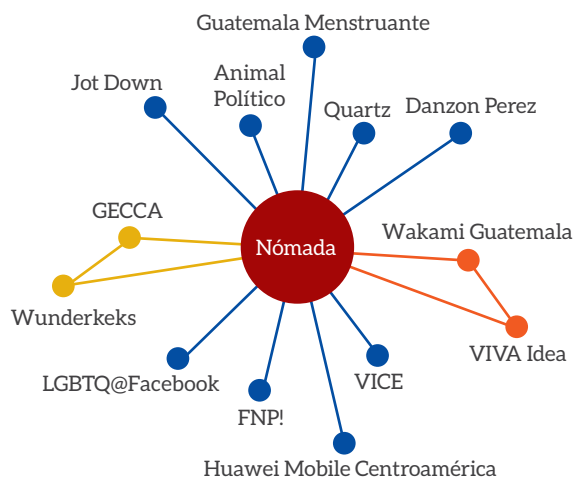
3. 1. Análise de Rede – Página Nómada

A partir da ferramenta *page like network*, do Netvizz, extraímos dados da página Nómada, no Facebook, nas duas profundidades possíveis: *depht 1*, que se refere à extração da rede formada por *likes* que a página Nómada deu ou recebeu diretamente; e *depth 2*, que observa não apenas os *likes* da página Nómada, mas também os *likes* das páginas que ela curtiu. Esse segundo processo constitui uma rede mais complexa, mostrando conexões indiretas que também constituem o nicho de interesse em que a página Nómada está circunscrita.

As figuras a seguir foram produzidas através do uso do aplicativo Gephi e são constituídas de páginas do Facebook (pontos) e suas conexões entre si (linhas). As páginas nomeadas são as que contam com o maior número de conexões.

Para uma melhor visualização, estabelecemos critérios de relevância principalmente em relação a dois quesitos: modularidade e grau (RECUERO, 2014). A modularidade é relacionada com a formação de bolhas devido a conexões das páginas entre si (evidenciada através das cores). Já o grau é relacionado ao número de conexões que uma página mantém com outras na rede analisada (o que é explicitado através do tamanho do ponto – quanto maior mais conexões tem a página).

[Figura 2]
Rede de conexões por likes da página Nómada em profundidade baixa (depht 1)



Fonte: Nómada (2018).

[Tabela 1]
Descrição das conexões por likes da página Nómada

Página	Categoria	Modularidade	Grau de entrada	Grau de saída	Grau
Nómada	Media/News Company	0	2	13	15
Wakami Guatemala	Retail Company	1	1	2	3
GECCA	Consulting Agency	2	1	2	3
VIVA Idea	Nonprofit Organization	1	2	0	2
Wunderkeks	Bakery	2	2	0	2
Huawei Mobile Centroamérica	Product/Service	0	1	0	1
FNPI	Nonprofit Organization	0	1	0	1
Animal Político	News & Media Website	0	1	0	1
Quartz	Media/News Company	0	1	0	1
VICE	Media/News Company	0	1	0	1
Jot Down	Magazine	0	1	0	1
LGBTQ@Facebook	Community	0	1	0	1
Guatemala Menstruante	Community	0	1	0	1
Danzon Perez	Community	0	1	0	1

Fonte: Nómada (2018).

Essa primeira extração nos permite visualizar uma rede com a página Nómada como central, ou seja, o nó de origem. Observamos que Nómada tem o número maior de conexões dessa rede, já que as outras páginas pouco fazem conexão entre si. Isso evidencia como há uma certa heterogeneidade aqui disposta. Mas, ao analisar que tipo de conexão é feita, observamos que a maior parte dos *likes* partiu da página Nómada (por exemplo, o grau de saída 13), estatística que dá a ver que essa rede é constituída em sua maioria por interação ativa da página Nómada.

No que toca ao jornalismo, as páginas Animal Político, Quartz, VICE e Jot Down e FNPI são as principais conexões. Observamos que a linha editorial desses veículos se assemelha muito em forma e conteúdo, flertando com o jornalismo cultural por uma via ideológica mais alinhada às políticas progressistas. A extração de dados evidencia que existem também conexões com páginas de caráter ativista relacionado a políticas de gênero e sexualidade, como a LGBT@Facebook e a Guatemala Menstruante.

Chegamos, pois, a algo pouco improvável: a rede de *likes* reflete basicamente a linha editorial do jornal Nómada. Mas a partir disso podemos justamente perceber o quanto a própria interface do Facebook influi na construção das redes. O like é um gostar. E as redes online, inclusive dos veículos de jornalismo, são constituídas através do gostar, da concordância, do interesse. É claro que se pensarmos do ponto de vista do jornalista, este constitui uma série de percursos que buscam contornar diversas barreiras, inclusive as algorítmicas, para a apuração de notícias.

No entanto, o usuário comum, o leitor, não necessariamente despende tempo tentando alcançar informações que estejam para além da bolha de interações das páginas que acompanha. É muito mais provável que um usuário acabe conhecendo uma página que esteja na rede de conexões da página que já acompanha. Nesse sentido, cabe no mínimo pensarmos o que significa a diversidade jornalística, onde ela começa e onde ela termina, e qual é o alcance do jornalismo independente. E, pelo que estamos vendo, não significa apenas em relação às notícias, mas as próprias interações das páginas podem ser importantes nesse quesito.

3. 2. Jota e Meio – Social Monitor

O que podemos ver a partir dos dados extraídos pela ferramenta Social Monitor,¹⁴ criada para orientar a compreensão de jornalistas e gestores que trabalham em veículos de mídia, dos assuntos mais propagados em tempo real, são as métricas geradas a partir dos dados públicos das publicações realizadas no Facebook e que podem ser consolidadas pela ferramenta ao longo de um determinado período, como por exemplo, no nosso caso, 19 dias. Assim, em cada dia, temos o número de posts de cada marca, o número dos likes, reações, comentários e compartilhamentos e, ainda, seguidores que a marca perdeu ou ganhou. São as métricas geralmente mais utilizadas para a análise da eficácia de propagação, pois geram números compreensíveis; por exemplo, no dia 20 de maio, dia posterior ao casamento da família real britânica, o post com maior índice de viralidade e com

[14] Dados acessados em 20 de maio de 2018.

mais de 19 mil likes foi da página “Ajudar o povo de humanas a fazer miçanga”, que postou a seguinte imagem:

[**Figura 3**]

**Post com índice mais alto de aceitação
no Facebook em 20 de maio de 2018**



maria capitolina
@antrindade

**eu poderia dormir hj até tarde
igual uma princesa disney
mas infelizmente meu relógio
biológico já acostumou a
ser empregado do sistema
capitalista**

Fonte: Social Monitor (2018)¹⁵

Diferente da análise anterior, da página Nómada, que nos fala da rede criada em torno e pela marca, aqui não conseguimos visualizar uma dinâmica de como a mensagem se propaga pela rede ou por quem ou quais mecanismos tornam possível que aquela mensagem se propague mais do que outras e em que nichos. O que conseguimos apreender nesse caso é somente o efeito que a mensagem nesta rede desconhecida aparece aos nossos olhos, ou seja, a audiência medida - as métricas de cada postagem. Anteriormente, os administradores das páginas podiam escolher abrir publicamente o nível que permitia a análise feita a partir da Nómada, mas, atualmente, estes dados foram fechados a priori pelo sistema do Facebook.

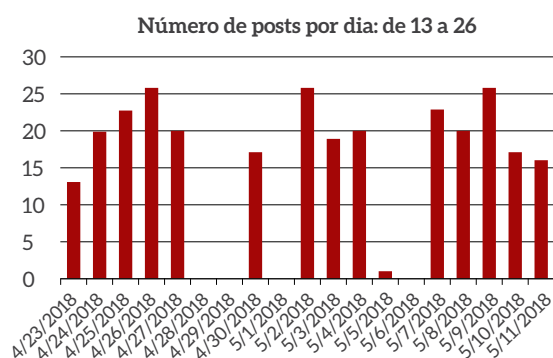
[15] SOCIAL MONITOR, c2018. Disponível em: <<https://bit.ly/2wVKaQJ>>. Acesso em: 20 maio de 2018.

Dessa forma, visualizamos o que é de interesse do público no Social Monitor. É importante citar também que os dados que nos permitiram visualizar e analisar a Nómada não estão abertos pelo Jota e Meio. No Social Monitor, conseguimos rapidamente saber, por exemplo, que o Jota posta mais de dez vezes por dia (ver figura 3) e que, no seu dia (24/4/2018) de mais compartilhamentos (1080), ganhou novos 58 seguidores, enquanto que em outro dia (25/4) de menos compartilhamentos (390), ganhou 102 novos seguidores. Pode-se analisar cada post de cada dia e cruzar os dados do conteúdo versus mais comentados, os que, por exemplo, geraram mais reações de raiva, os mais comentados, ler os comentários na página da marca, e assim por diante. Mas, fora a análise de conteúdo e das métricas de propagação, que interessam diretamente aos estrategistas de conteúdo das marcas, o que o público em geral pode esperar desse tipo de extração e possibilidade de análise de dados é o efeito e as discussões que as mensagens (*posts*) geram. Isso dá uma ideia do que repercute na esfera pública criada pelo FB. No entanto, o quanto essa esfera é ou não pública (no sentido de que nem todos que interagem recebem a mesma informação no seu *feed* de conteúdo e, ainda, que o alcance das páginas foi reduzido em comparação com o alcance dos posts de amigos e da família) seria outra discussão para a qual não teremos mais espaço neste artigo, mas a qual endereçamos para a discussão da propagação do jornalismo independente na mídia social.

O Meio atua basicamente como uma *newsletter* diária enviada diretamente ao público cadastrado e adotou o Facebook mais como uma estratégia de pre-

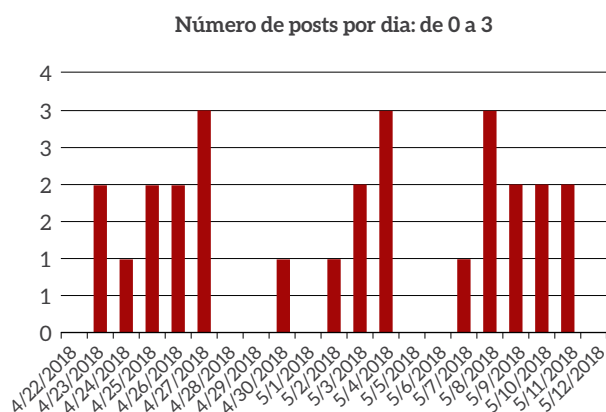
sença, e não de propagação de conteúdo. O Meio posta no máximo duas vezes por dia (ver figura 5), e os posts são linkados para o texto da *newsletter*, disponível em um site. Através da comparação entre as estatísticas das duas marcas geradas por essa ferramenta, vemos que a quantidade de posts por dia influencia no número de likes e compartilhamentos, bem como na frequência com que cada marca ganha novos seguidores. Com 10 a 20 posts por dia, o Jota ganha novos seguidores frequentemente, enquanto o Meio tem um índice de perda frequente.

[Figura 4]
Número de posts do Jota por dia.



Fonte: Social Monitor (2018).

[Figura 5]
Frequência de postagem do Meio



Fonte: Social Monitor (2018).

Considerações Finais

As duas análises implicam na visão de duas dinâmicas diferentes das redes, que não são excludentes, ao contrário, pois poderiam ser sobrepostas para que melhor pudéssemos entender tanto o próprio funcionamento do modelo de propagação de conteúdo do Facebook, ou seja, seu algoritmo, quanto a eficácia e o alcance do jornalismo independente. No entanto, não foi possível aplicar a mesma metodologia de análise para todas as marcas, de forma a sobrepor as análises. O Social Monitor, por outro lado, serve às marcas brasileiras, e não analisa marcas estrangeiras, como o caso da Nómada.

A modularidade, relacionada com a formação de bolhas devido a conexões das páginas entre si, e o grau, relacionado ao número de conexões que uma página mantém com outras na rede analisada, são os atributos dos dados que nos possibilitam a visualização de territórios de sentidos gerados a partir da interação na rede social do Facebook. A possibilidade de observação da circulação dos conteúdos nestes territórios, o que poderíamos nomear de formação de uma esfera pública temporária, no caso do Meio e Jota, foi impossível de ser entendida no nosso caso, pelas limitações expostas no artigo.

Assim, os dados circulam, mas não são puros; são mediados a partir de modelos cujo funcionamento desconhecemos. Para ter um acesso mínimo a esses dados, mesmo que mediados, o pesquisador da área da comunicação precisa desenvolver meios de acesso a esses dados de forma interdisciplinar e quase sempre envol-

vendo um grupo, como é o caso de dois em destaque no cenário brasileiro. Como o Laboratório de Estudos sobre Imagem e Cibercultura (LABIC - <http://www.labic.net/>), locado na Universidade Federal do Espírito Santo, comandado pelo docente Fábio Malini e que atua em parceria com o Programa de Pós-Graduação em Ciência da Computação daquela instituição. Na Universidade de São Paulo, temos o docente Pablo Ortellado coordenando o grupo de pesquisa Políticas Públicas para o Acesso à Informação (Gpopai), ligado ao curso de Gestão Pública na USP Leste.

Ou seja, o futuro da pesquisa em comunicação digital nos parece apontar para a interdisciplinariedade, para a formação de redes entre pesquisadores e para uma clareza sobre a natureza dos dados extraídos das diversas redes sociais que operam com base no *input* e cruzamento de dados que formam, hoje, o panorama da comunicação social contemporânea. Esse processo pode trazer mais transparência na compreensão da natureza dos dados gerados pelos diversos formatos de rede social e, consequentemente, mais transparência no acesso às informações e fortalecimento da dimensão democrática da comunicação digital. ■

[DANIELA OSVALD RAMOS]

Professora e pesquisadora no Departamento de Comunicações e Artes da Escola de Comunicações e Artes da Universidade de São Paulo e no Programa de Pós-Graduação em Ciências da Comunicação (PPGCOM). Pesquisadora associada ao Núcleo de Estudos da Violência (NEV-USP). É integrante dos grupos de pesquisa: COM+ (ECA-USP) e Semiótica da Comunicação.

E-mail: dosvald@gmail.com

[EGLE MÜLLER SPINELLI]

Doutora em Ciências da Comunicação (ECA-USP). Professora e pesquisadora do curso de Jornalismo e do Programa de Pós Graduação em Comunicação e Práticas de Consumo, da Escola Superior de Propaganda e Marketing (PPGCOM-ESPM). É integrante dos grupos de pesquisa: COM+ (ECA-USP) e Lógicas e Modelos de Gestão em Jornalismo (ESPM-SP).

E-mail: egle.spinelli@espm.br

[MÁRIO ARRUDA]

Doutorando no Programa de Comunicação e Informação da Universidade Federal do Rio Grande do Sul. Membro do Grupo de Pesquisa Semiótica e Culturas da Comunicação (GPESC).

E-mail: marioarruds@gmail.com

Referências

ALVES, Marcelo. Abordagens da coleta de dados nas mídias sociais. In: SILVA, Tarcízio; STABILE, Max (Orgs.). **Monitoramento e pesquisa em mídias sociais**: metodologias, aplicações e inovações. São Paulo: Uva Limão, 2016.

ARRUDA, Mario. **Ecologia da bolha algorítmica**: liberdade e controle nas redes de comunicação online. 2018. 147 f. Dissertação (Mestrado em Comunicação e Informação) – Universidade Federal do Rio Grande do Sul, Porto Alegre, 2018.

AVISO importante! Facebook muda API para visualização de informações dos usuários. **Buzzmonitor**, 16 fev. 2018. Disponível em: <<https://bit.ly/2BCKPHp>>. Acesso em: 20 abr. 2018.

CANAL MEIO S.A. **Meio**, 2016-2018. Apresenta notícias do Brasil e do mundo. Disponível em: <<https://www.canalmeio.com.br/>>. Acesso em: 8 jan. 2018.

CONSTINE. Josh. Facebook restricts APIs, axes old Instagram platform amidst scandals. **TechCrunch**, San Francisco, 4 abr. 2018. Disponível em: <<https://tcrn.ch/2M3PZ5b> />. Acesso em: 20 abr. 2018.

FRAGOSO, Suely; RECUERO, Raquel; AMARAL, Adriana. **Métodos de pesquisa para internet**. Porto Alegre: Editora Sulina, 2011.

GRAPH API. In: Facebook, 2018. Disponível em: < <https://bit.ly/1mAJRfT> >. Acesso em: 29 mar. 2018.

JOTA: notícias que fazem diferença, 2014. Apresenta notícias sobre os poderes executivo, legislativo e judiciário. Disponível em: <<https://www.jota.info/>>. Acesso em: 6 jan. 2018.

NÓMADA, c2014. Apresenta notícias sobre política e cotidiano. Disponível em: <<https://nomada.gt/>>. Acesso em: 22 fev. 2018.

O'NEIL, Cathy. **Weapons of math destruction**: how big data increases inequality and threatens democracy. New York: Broadway Books, 2016.

POULPIQUET, Pierre de. What is a Walled Garden? And why it is the strategy of Google, Facebook and Amazon Ads platform? **Medium**, 3 nov. 2017. Disponível em: <<https://bit.ly/2IOcXYL>>. Acesso em: 19 maio 2018.

RECUERO, Raquel; BASTOS, Marco; ZAGO, Gabriela. **Análise de redes para mídia social**. Editora Sulina, Porto Alegre, 2015.

RECUERO, Raquel. Contribuições da análise de redes sociais para o estudo das redes sociais na internet: O caso da hashtag #Tamojuntodilma e #CalaabocaDilma. **Revista Fronteiras**: estudos midiáticos, São Leopoldo, v. 16, n. 2, p. 1-18, maio/ago. 2014. Disponível em: <<https://bit.ly/2LnGDhe>>. Acesso em: 20 abr. 2018.

SOCIAL MONITOR. Disponível em: <<http://www.socialmonitor.com.br/>>. Acesso em: 23 mar. 2018.

THE GEPHI CONSORTIUM. Gephi: makes graphs handy, 2008-2017. Disponível em: <<https://gephi.org/>>. Acesso em: 28 mar. 2018.