



# XVI Congreso de la Asociación Latinoamericana de Investigadores de la Comunicación (ALAIC)

*La Comunicación como Bien Público Global:*

*Nuevos lenguajes críticos y debates hacia el porvenir*

**Buenos Aires, Argentina, 26 al 30 de septiembre de 2022**

Organizan

- ❖ Asociación Latinoamericana de Investigadores de la Comunicación (ALAIC).
- ❖ Federación Argentina de Carreras de Comunicación Social (FADECCOS).

## Ponencia presentada al GT 19

### Impactos da ferramenta Twitter Monitor para pesquisas acadêmicas sobre redes sociais digitais.

Impacts of the Twitter Monitor tool for academic research on digital social networks.

Josir Cardoso Gomes <sup>1</sup>

Ricardo Medeiros Pimenta <sup>2</sup>

**Resumen:** Com expressivo alcance global, o Twitter é uma rede social digital com grande repercussão no Brasil e por conta deste alcance e por ter uma interface aberta, é muito utilizado na pesquisa acadêmica na pesquisa de política, sociologia e comunicação social buscam trazer novos conhecimentos sobre a ambiência que circula por tal rede. Este trabalho relata como o uso da ferramenta Monitor Twitter foi idealizada pelo Laboratório em Rede de Humanidades Digitais (LARHUD) e como tem sido utilizada por pesquisadores que gostariam de analisar os dados do Twitter em suas pesquisas. Utilizando o método de pesquisa-ação, buscou-se responder às seguintes questões: em que medida o software

---

<sup>1</sup> Josir Cardoso Gomes. IBICT/UFRJ, Doutor em Ciência da Informação. Brasil. josir@irdx.com.br

<sup>2</sup> Ricardo Medeiros Pimenta. IBICT. Doutor em Memória Social. Brasil. ricardopimenta@ibict.br



auxiliou na pesquisa de seus usuários e quais os conhecimentos adquiridos pelos autores ao prover a democratização do acesso a este serviço para a comunidade acadêmica? Como resultado, de um total de 9 projetos iniciados, 4 já foram concluídos gerando um artigo em periódico, um capítulo de livro e uma dissertação de mestrado já qualificada. Foi criado também um processo de acompanhamento dos projetos pois percebeu-se que não bastava fornecer a tecnologia: era necessário auxiliar as pesquisadoras e pesquisadores na definição de critérios de busca e para filtrar registros que não deveriam ser incorporados no projeto de pesquisa.

**Palabras Clave: Humanidades Digitais, Redes Sociotécnicas, Twitter, Pesquisa Ação**

**Abstract:** With significant global reach, Twitter is a digital social network with great repercussion in Brazil and because of this reach and for having an open interface, it is widely used in academic research in politics, sociology and social communication seeking to bring new knowledge about the ambience that circulates through such network. This paper reports on how the use of the Twitter Monitor tool was devised by the Networked Laboratory of Digital Humanities (LARHUD) and how it has been used by researchers who would like to analyse Twitter data in their research. Using the action-research method, we sought to answer the following questions: to what extent the software has helped the research of its users and what is the knowledge acquired by the authors by providing the democratization of access to this service for the academic community? As a result, from a total of 9 initiated projects, 4 have already been concluded, generating an article in a journal, a book chapter and a qualified master's dissertation. A process of monitoring the projects was also created because it was realized that it was not enough to provide the technology: it was necessary to help the researchers in the definition of search criteria and to filter records that should not be incorporated into the research project.

**Key words: Digital Humanities, Social Networks, Twitter, Action Research**



## Introdução

Com expressivo alcance global, ainda que não lidere o universo das redes sociotécnicas, o Twitter é uma rede social digital com enorme repercussão no Brasil, já que é um canal muito utilizado pelos veículos jornalísticos de grande porte, pela mídia alternativa, por políticos, artistas e por uma boa parte dos chamados influenciadores digitais. Segundo o relatório da DataReportal (DataReportal, 2022), o Twitter contava com 19,05 milhões de usuários ou 10,8% dos usuários de Internet no Brasil até o início de 2022. Por conta deste alcance e por ter uma interface aberta, a rede Twitter é muito utilizada na pesquisa acadêmica na área das humanidades em geral. Pesquisas nas áreas de ciência política, sociologia e comunicação social buscam trazer novos conhecimentos sobre a ambiência que circula por tal rede. Diferentemente do Facebook, Instagram, Whatsapp e TikTok que não têm uma interface computacional pública onde se possa coletar os dados gratuitamente, a rede Twitter permite a coleta a partir de termos de busca similares ao formato de busca já conhecido pelos usuários da plataforma.

Esta interface pública é denominada API (Application Programming Interface) e permite que dois sistemas computacionais se comuniquem através de um protocolo pré-estabelecido. Apesar da API ser aberta e gratuita, a coleta de dados não é trivial pois requer primeiramente uma autorização formal da empresa que detém o controle sobre a



plataforma e, uma vez de posse desta autorização, o processo também demanda uma competência computacional que grande parte dos pesquisadores das humanidades e ciências sociais aplicadas não detêm.

Desta forma, a partir da necessidade de vários pesquisadores que gostariam de utilizar os dados do Twitter em suas pesquisas mas não tinham possibilidade à API, o Laboratório em Rede de Humanidades Digitais (LARHUD)<sup>3</sup> do PPGCI/IBICT<sup>4</sup> construiu a ferramenta Monitor Twitter. Este artigo traz o relato da construção e uso do software que visa coletar dados do Twitter para usuários que não detêm as condições técnicas e computacionais para tal intento. Utilizando o método de pesquisa-ação, buscou-se responder às seguintes questões: em que medida o software auxiliou na pesquisa de seus usuários e quais os conhecimentos adquiridos pelos autores ao prover a democratização do acesso a este serviço para a comunidade acadêmica?

## Referencial Teórico

A plataforma Twitter é controlada pela empresa norte-americana de capital aberto Twitter Inc. Em 2021, a corporação obteve um faturamento global de 5,08 bilhões de dólares (Twitter Inc, 2021) e diariamente uma média de 300 milhões de usuários acessam

<sup>3</sup> <http://www.larhud.ibict.br/>

<sup>4</sup> <http://www.ppgci.ufrj.br/pt/>



a plataforma (GDELT Project, 2022). Apesar de ter um faturamento bem inferior ao das empresas que compõem a GMAFIA (Google, Microsoft, Apple, Facebook, IBM e Amazon), a corporação Twitter segue o mesmo modelo econômico baseado na venda de anúncios publicitários. Se excluirmos a venda direta de produtos físicos (que compõem a maior parte do faturamento da Amazon e da Apple), venda de licenças de uso de software (Microsoft e IBM), a construção de plataformas sócio-digitais (Dantas, 2019) a partir do trabalho digital alienado (Fuchs & Sandoval, 2014) e da da comercialização de publicidade através destas plataformas se tornou um excelente negócio.

Do outro lado do mundo, as chinesas Baidu, Alibaba, Tencent e Byte Dance seguem o mesmo caminho e o TikTok passou a ser a primeira plataforma concorrente de peso no mercado ocidental das redes sócio-digitais norte-americanas (Abidin, 2020). Entretanto, a concentração de acesso continua sendo das grandes redes sócio-digitais americanas: Facebook, Instagram, Youtube e Twitter. O WhatsApp não foi incluído aqui por não existir unanimidade de que a plataforma seja considerada como uma rede sócio-digital já que não existe nenhum tipo de conversação pública, ou seja, todo o tráfego da plataforma opera em canais privados.

Em linhas gerais, a plataforma Twitter permite que o usuário execute 4 ações primariamente: (1) que ele envie textos, comentários, links ou avisos sobre qualquer



assunto publicamente. Cada texto enviado é denominado “tweet”. Nesse momento ele é o produtor de uma informação para a rede; (2) que ele siga outros usuários, isto é, ele indica que quer ler os tweets de outro usuário; (3) que ele possa curtir o tweet de outro usuário e, da mesma forma, o seu tweet pode receber curtidas de outros usuários e, por fim, (4) ele pode “retweetar”, ou seja, repassar para a sua rede um tweet de outrem. Tal como descreve Lúcia Santaella e Renata Lemos,

Quando seguimos alguém no Twitter, estamos fazendo uma “assinatura” do seu canal de informações. Desta forma, cada pessoa que seguimos torna-se um canal provedor de um determinado tipo de conteúdo para nosso fluxo pessoal, agregando o seu conteúdo ao total de nosso inflow. (Santaella & Lemos, 2010, p.73)

O conjunto destas ações fará com que a ferramenta crie uma linha de tempo que contenha uma lista dos tweets recomendados pelo algoritmo do Twitter que irá aparecer sempre que um usuário entra na plataforma.

Como na maioria das redes sócio digitais, as regras do algoritmo são opacas, ou seja, o usuário não define ou não sabe qual é o critério utilizado para que um tweet apareça na sua timeline. O usuário, de certa forma, auxilia o algoritmo nessa seleção, quando se conecta a outros usuários ou quando curte tweets sobre um determinado assunto mas, no frígir dos ovos, quem define o que o usuário irá visualizar será o algoritmo.



Sob o ponto de vista de abertura de dados, o Twitter se diferencia das demais redes pois permite que usuários regulares possam capturar os dados da plataforma através de uma API. O Facebook também oferecia um serviço semelhante, mas com o escândalo da Cambridge Analytica (Schneble et al., 2018), a corporação bloqueou o acesso às buscas através de sua API (Venturini & Rogers, 2019).

Apesar de existirem critérios bem documentados no site do Twitter, a autorização para o uso da API é dada ao bel prazer da empresa, ou seja, não basta se adequar aos critérios indicados pela empresa, existem fatores não explicitados pela empresa para permitir ou não o acesso. Ou seja, mesmo que você seja um pesquisador reconhecido, não existe garantia de que a empresa irá lhe conceder acesso aos dados para a sua pesquisa.

Uma vez que a autorização é concedida, a corporação fornece uma “chave de API” e, a partir desta chave, é possível coletar os dados a partir de diversos serviços oferecidos pela plataforma. É possível, por exemplo, obter os tweets a partir de um critério de busca ou os tweets de um determinado perfil. Essa consulta, entretanto, deve ser realizada através de um programa de computador, ou seja, o usuário precisa saber programar em alguma linguagem computacional para que possa utilizar este serviço.

Existem diversas ferramentas externas que permitem o uso da API sem a necessidade de programação, entretanto elas próprias também são opacas, ou seja, não



informam como a coleta é realizada ou se os dados que você está coletando serão utilizados por terceiros. Em função da limitação financeira dos laboratórios no Brasil, as dezenas de ferramentas pagas disponíveis não foram sequer avaliadas. Yu e Muñoz-Justicia avaliaram diversas opções gratuitas e mostraram que algumas requerem que sejam instalados softwares proprietários enquanto outras são online mas não tem o código fonte aberto (Yu & Muñoz-Justicia, 2022). São exemplos de ferramentas gratuitas, a TAGS, Mozdeh e Netlytic.

Um dos pontos centrais que devem ser analisados é que a API fornece apenas uma amostra dos tweets que atendam ao critério definido pelo usuário. Ou seja, o algoritmo do que é relevante também atua na API e não apenas na linha de tempo que é visível para o usuário. Existe a possibilidade de obter todos os tweets mas somente para um serviço específico que é pago denominado Firehose.

Percebe-se assim que a API também é uma fonte de receita para o Twitter. Dos 5 bilhões faturados em 2021, 571 milhões vieram da disponibilização de dados para terceiros. Além da API gratuita, a empresa disponibiliza a API Enterprise que permite a coleta de um volume maior que a versão gratuita e também a carga via Firehose. Posto isso, pode-se supor que a empresa deve manter o serviço de API como forma de fomentar o seu uso.



## O uso de dados do Twitter em Pesquisas Acadêmicas

Devido ao seu alcance midiático e em função da democratização de sua API, os dados do Twitter são muito utilizados em pesquisas acadêmicas. Pode-se dividir tais pesquisas em duas categorias: a análise sociológica da rede em si e o uso dos fluxos comunicacionais gerados nela como suporte para análise de discursos e padrões de comportamento.

Na primeira categoria, pode-se citar o estudo sobre desinformação, sobre o processo de midiatização e sobre o comportamento dos usuários em função de eventos e polêmicas que surgem especificamente nesta rede e transbordam para a “vida real”. No estudo sobre fenômenos de desinformação, pode-se citar como exemplo o trabalho de Bastos que estudou o volume de contas que desapareceram após o Brexit (Bastos, 2021), demonstrando assim como o uso de robôs (bots) buscou influenciar o debate político. Já no Brasil, a análise da atuação de robôs nas eleições presidenciais de 2018 também suscitou pesquisas interessantes (Paula & Michalski, 2019).

Já a análise das redes que se formam entre a interação entre os perfis do Twitter também enseja diversas análises. A partir deste tipo de análise, é possível encontrar grupos (ou *clusters* no jargão da mineração de dados) tal como o mostrado na figura abaixo:

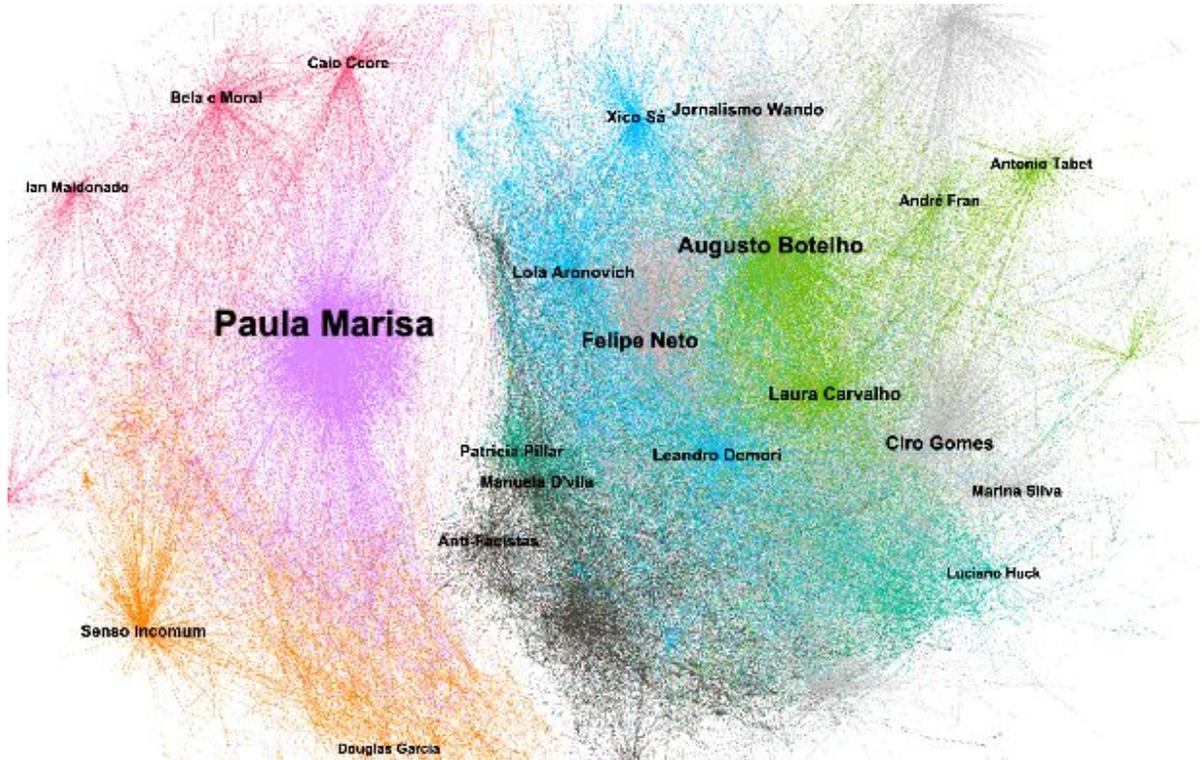


Figura 1: Repercussão da Rede Sleeping Giant. Fonte: Oliveira, 2022

Para a segunda categoria, os dados obtidos no Twitter servem como apoio quantitativo e qualitativo para pesquisas que tratem sobre um determinado tema que é discutido na rede. Aqui entram pesquisas na área de análise de sentimentos e processamento de linguagem natural (Cury, 2019) e que buscam analisar os tweets de determinadas figuras públicas e seus apoiadores utilizando métodos baseados na análise do discurso (Neto, 2019).

## Resultados



O projeto de construção do Monitor Twitter se deu a partir do interesse de diversos pesquisadores em coletar dados do Twitter que vinham ao laboratório perguntar como essa coleta poderia ser realizada. A partir desta demanda, buscou-se alguma ferramenta que atendesse aos nossos requisitos de volume de coleta e também de ser uma ferramenta aberta e não-proprietária já que o PPGCI/IBICT sempre busca ferramentas que se alinhem com os princípios da ciência aberta.

A única ferramenta encontrada à época que atendia aos requisitos para captura de dados e que vinha com o código aberto foi a DMI-TCAT<sup>5</sup> desenvolvida pela Digital Methods Initiative (DMI) formada por diversos grupos de pesquisa europeus. Entretanto, ela foi descartada pois não haviam desenvolvedores no LARHUD aptos a implementar a ferramenta desenvolvida pelos europeus.

Desta forma, optou-se por construir uma ferramenta própria que auxiliasse na captura dos dados do Twitter e que a programação da coleta pudesse ser realizada por pesquisadores que não tivessem conhecimento de programação ou de TI. O desenvolvimento da primeira versão foi realizado por 3 programadores e demorou 4 meses para ser concluído.

---

<sup>5</sup> <https://wiki.digitalmethods.net/Dmi/ToolDmiTcat>



Após a construção da primeira versão do software, comunicamos a disponibilidade da ferramenta para diversos pesquisadores de diversos programas de pós-graduação na área da Comunicação, Ciências Sociais e Ciência da Informação através de palestras e eventos e também criamos uma página no laboratório<sup>6</sup> explicando o que fazia o software e também sobre o caráter gratuito e experimental do mesmo.

Para cada pesquisador que procurava o laboratório, criava-se um projeto que incorporava os termos de busca e os resultados coletados em tempo real e, a partir daí, o pesquisador podia obter os dados coletados para uso em planilhas eletrônicas no formato 'csv' ou para uso no software Gephi (Bastian et al., 2009) no formato 'gfx'. Com o resultado em mãos, o pesquisador poderia solicitar a inclusão de novos termos ou refinar a busca para atingir um resultado que se adequasse à sua pesquisa. Na figura 2, é apresentada a tela do resultado de um projeto no Monitor Twitter.

---

<sup>6</sup> <http://monitor.farmi.pro.br/>



## Livros Amazon

### Termos:

amazon livro algoritmo  
amazon algoritmo livros  
amazon livro  
algoritmo amazon

Total de Tweets Coletados: 4.493

Total de Retweets: 5.777

Total de Usuários Únicos Coletados: 2.766

Alcance Total: 14814

### Top 3 - Tweets mais favoritos:

1. [Link](#): absolutely horrifying: amazon's algorithm identified a product that people were using to commit suicide. it then started to recommend additional products that r
2. [Link](#): El algoritmo de Amazon identificó un producto que la gente usaba para suicidarse. Y detectó que la gente lo compraba con otros que hacían el suicidio más fácil. <https://t.co/htLuRy8yY>
3. [Link](#): Vendo aqui que o pessoal já tá recebendo email da Amazon com a data estimada de entrega do livro, já chegou esse email pra vcs moots? Fiquem ligados!

### Palavras mais utilizadas:

posso - 530  
indicar - 529  
leve - 482  
oferta - 454  
leitura - 358  
#oferta - 333  
comprar - 317  
#desconto - 315  
#promoção - 315  
#cupom - 311  
espero - 295  
#booktwitter - 290  
#bookstan - 288  
ajude - 287  
ressaca - 283  
matador - 283  
risada - 283  
oferta: - 282

Tweets por faixa de horário

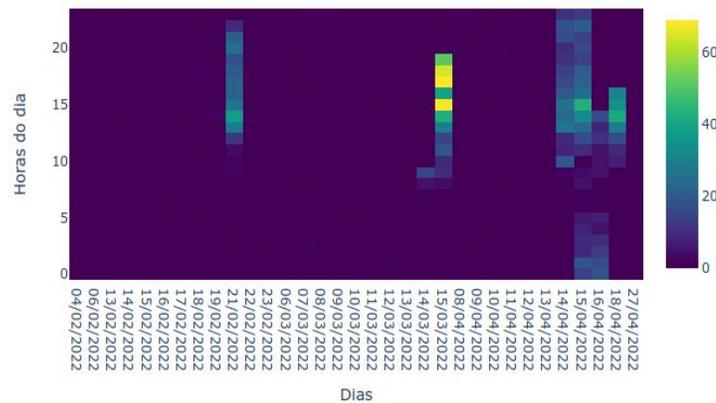


Figura 2: Estatística de um projeto no Monitor Twitter (Fonte: o Autor)



O laboratório, por sua vez, consegue visualizar todos os projetos em andamento e como está a carga de cada um deles, tal como pode se observar na figura 3.

Seleccione projeto para modificar ADICIONAR PROYECTO + EXPORTACIÓN GEPHI

Ação:   0 de 8 seleccionados

<input type="checkbox"/>	NOME	USUARIO	STATUS	TOT TWTTS
<input type="checkbox"/>	Influenciadores	admin	Concluído	2706
<input type="checkbox"/>	Educação 2019	admin	Concluído	55965
<input type="checkbox"/>	IBICT	admin	Ativo	258
<input type="checkbox"/>	CNPq	admin	Concluído	16269
<input type="checkbox"/>	Wikileaks	tiago	Concluído	14529
<input type="checkbox"/>	Medical Physics	admin	Concluído	1951
<input type="checkbox"/>	Primeiro Assédio	admin	Processando	196832
<input type="checkbox"/>	Humanidades Digitais	admin	Ativo	15130

8 projetos

Figura 3: Projetos em andamento (Fonte: o Autor)

De um total de 9 projetos, 4 já foram concluídos e 5 ainda estão em andamento. Os dados dos projetos concluídos foram utilizados em um artigo acadêmico, um capítulo de livro e uma dissertação de mestrado já qualificada. Até o momento foram coletados mais de 3,3 milhões de tweets e 3,5 milhões de usuários únicos e pretende-se também incorporar os *datasets* públicos fornecidos pelo próprio Twitter ou por outros pesquisadores para que se possa realizar pesquisas em eventos passados.

Ao longo das coletas e das entrevistas com os pesquisadores, percebeu-se que o software em si não bastava para auxiliar os pesquisadores e foi necessário criar um processo de aprendizagem e acompanhamento dos projetos, ou seja, não bastava



fornecer a tecnologia: era necessário acompanhar e auxiliar o pesquisador para que este pudesse incluir os critérios de busca que o Twitter esperava e para filtrar registros que não deveriam ser incorporados no projeto de pesquisa. Além disso, diversas melhorias foram sugeridas pelos pesquisadores que foram incorporadas ao software.

Ao longo da coleta, foram encontrados alguns erros e problemas de escalabilidade já que a quantidade de tweets coletados ultrapassaram algumas vezes a casa de centenas de milhares de tweets e quando muitas pesquisas estavam ativas, era necessário escalonar manualmente quais as pesquisas teriam precedência de execução já que a API tem limite de número de tweets a recuperar em um dado intervalo.

Havia também a preocupação de não disponibilizar uma quantidade de registros que ultrapassasse o limite imposto pelos termos de uso do Twitter. O Twitter indica em seu termo de uso que “Pesquisadores acadêmicos terão permissão para distribuir um número ilimitado de IDs de Tweet e/ou IDs de Usuário se estiverem representando uma instituição acadêmica e somente para pesquisa não comercial.”<sup>7</sup>, entretanto o pesquisador necessita do texto do tweet e nesse caso só poderia ser enviado até 50.000 diários.

---

<sup>7</sup> <https://developer.twitter.com/pt/developer-terms/agreement-and-policy>



## Conclusões

A partir do relato das diversas iniciativas bem sucedidas e também dos problemas encontrados, o resultado superou as expectativas já que a ferramenta conseguiu fornecer os dados necessários para que os usuários pudessem realizar suas investigações acadêmicas de maneira colaborativa. Ficou claro que a implementação da ferramenta Monitor Twitter foi bem sucedida. Sem o software, algumas pesquisas não poderiam ter sido realizadas já que os pesquisadores não teriam a competência computacional necessária para coletar os dados ou se conseguissem utilizar os softwares proprietários disponíveis não teriam o volume de dados suficiente para suas análises.

As solicitações e dúvidas dos pesquisadores também permitiu que a equipe do Larhud tivesse a oportunidade de entender os problemas no processo e incorporar melhorias ao software e também na forma de disponibilizar os dados para seus usuários.

A construção do software e o processo de aprendizado segue em fluxo contínuo, ou seja, novas funcionalidades continuam sendo incorporadas como novos relatórios e gráficos, rotinas para estimar contas-robôs (bots) e também a incorporação de rotinas de análise de sentimentos. Por fim, é importante ressaltar que o código-fonte está disponível



no Github<sup>8</sup> con una licencia abierta e cualquier institución puede reproducir o ambiente computacional necesario em seus próprios servidores.

## Bibliografía

- Abidin, C. (2020). Mapping Internet Celebrity on TikTok: Exploring Attention Economies and Visibility Labours. *Cultural Science Journal*, 12(1), 77–103. <https://doi.org/10.5334/csci.140>
- Bastian, M., Heymann, S., & Jacomy, M. (2009). Gephi: An open source software for exploring and manipulating networks. *International AAAI Conference on Weblogs and Social Media*. <https://gephi.org/>
- Bastos, M. (2021). This Account Doesn't Exist: Tweet Decay and the Politics of Deletion in the Brexit Debate. *American Behavioral Scientist*, 65(5), 757–773. <https://doi.org/10.1177/0002764221989772>
- Cury, R. M. (2019). Oscillation of tweet sentiments in the election of João Doria Jr. For Mayor. *Journal of Big Data*, 6(1), 42. <https://doi.org/10.1186/s40537-019-0208-1>
- Dantas, M. (2019). The Financial Logic of Internet Platforms: The Turnover Time of Money at the Limit of Zero. *TripleC: Communication, Capitalism & Critique. Open Access Journal for a Global Sustainable Information Society*, 17(1), 132–158. <https://doi.org/10.31269/triplec.v17i1.1088>
- DataReportal. (2022). *Digital 2022: Brazil*. DataReportal – Global Digital Insights. <https://datareportal.com/reports/digital-2022-brazil>
- Fuchs, C., & Sandoval, M. (2014). Digital Workers of the World Unite! A Framework for Critically Theorising and Analysing Digital Labour. *tripleC: Communication, Capitalism & Critique. Open Access Journal for a Global Sustainable Information Society*, 12(2). <https://doi.org/10.31269/triplec.v12i2.549>

---

<sup>8</sup> <https://github.com/josircg/twitsearch>



GDELT Project. (2022). *Visualizing A Decade Of Twitter's Evolution: Jan 2012 – Mar 2022 – The GDELT Project*. <https://blog.gdelproject.org/visualizing-a-decade-of-twiters-evolution-jan-2012-mar-2022/>

Neto, L. V. P. (2019). *Jair Bolsonaro e a busca pelo conflito permanente: Análise Crítica do Discurso das publicações no Twitter de Jair Bolsonaro durante a eleição presidencial brasileira de 2018* [Dissertação de Mestrado, Instituto Universitário de Lisboa]. [https://repositorio.iscte-iul.pt/bitstream/10071/19038/1/master\\_livino\\_pinheiro\\_netto.pdf](https://repositorio.iscte-iul.pt/bitstream/10071/19038/1/master_livino_pinheiro_netto.pdf)

Oliveira, D. C. (2022). A politização do consumo mediada pela tecnologia: O caso do Sleeping Giants no Brasil. Em *No prelo*.

Paula, L. T. de, & Michalski, R. (2019). Os bots de disseminação de informação na conjuntura das campanhas presidenciais de 2018 no Brasil. *Múltiplos Olhares em Ciência da Informação - ISSN 2237-6658*, 9(1). <http://portaldeperiodicos.eci.ufmg.br/index.php/moci/article/view/4066>

Santaella, L., & Lemos, R. (2010). *Redes sociais digitais: A cognição conectiva do Twitter* (1a ed). Paulus.

Schneble, C. O., Elger, B. S., & Shaw, D. (2018). The Cambridge Analytica affair and Internet-mediated research. *EMBO Reports*, 19(8). <https://doi.org/10.15252/embr.201846579>

Twitter Inc. (2021). *Fiscal Year Annual Report—2021—Twitter*. Twitter Inc. <https://investor.twitterinc.com/financial-information/annual-reports/default.aspx>

Venturini, T., & Rogers, R. (2019). “API-Based Research” or How can Digital Sociology and Journalism Studies Learn from the Facebook and Cambridge Analytica Data Breach. *Digital Journalism*, 7(4), 532–540. <https://doi.org/10.1080/21670811.2019.1591927>

Yu, J., & Muñoz-Justicia, J. (2022). Free and Low-Cost Twitter Research Software Tools for Social Science. *Social Science Computer Review*, 40(1), 124–149. <https://doi.org/10.1177/0894439320904318>