WhatsApp and transparency: an analysis on the effects of digital platforms' opacity in political communication research agendas in Brazil

Viktor Chagas; Gabriella Da-Costa

Nota: Este artículo se puede leer en español en:
https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/87120

Recommended citation:

Chagas, Viktor; **Da-Costa, Gabriella** (2023). "*WhatsApp* and transparency: an analysis on the effects of digital platforms' opacity in political communication research agendas in Brazil". *Profesional de la información*, v. 32, n. 2, e320223.

https://doi.org/10.3145/epi.2023.mar.23

Manuscript received on September 20th 2022
Accepted on February 1st 2023

Viktor Chagas ✉
https://orcid.org/0000-0002-1806-6062

Fluminense Federal University
Department of Media and Cultural Studies
Rua Professor Marcos Waldemar de Freitas Reis, s/n
Niterói, RJ CEP: 24210-201 Brazil
viktor@midia.uff.br

Gabriella Da-Costa
https://orcid.org/0000-0002-4234-4544

Fluminense Federal University
Communication Graduate Program
Rua Professor Marcos Waldemar de Freitas Reis, s/n
Niterói, RJ CEP: 24210-201 Brazil
gabrielladacosta@gmail.com

Abstract

This article aims to discuss what we call environmental opacity, a condition of mobile instant messaging services (MIMS) that operates on the basis of end-to-end encryption systems. Utilizing *WhatsApp* as a specific example, the article presents two fundamental dilemmas around which some issues concerning transparency are mobilized when it comes to digital private communication. The first of them relates to how end-to-end encryption has simultaneously become an asset and a problem for democratic environments; on the one hand, protecting users' privacy, and on the other, allowing for the circulation of misinformation and harmful content. The second dilemma deals with how this environment of opacity impacts the ethics and transparency of scholarly research focused on *WhatsApp* and other MIMSs. The paper also reviews an extensive body of studies that discuss the political uses of *WhatsApp* in different dimensions, and argues that emerging countries with large user bases, such as Brazil and India, have experienced a series of negative effects after the adoption of *WhatsApp* by politically oriented groups. Among the main proposals, the article suggests some measures to foster platform transparency and facilitate scientific research instead of hindering it.

Keywords

Algorithmic transparency; *WhatsApp*; Environmental opacity; Political communication; Privacy; Mobile instant messaging services; Research ethics; Policies.

Funding

This article benefits from the support of the *Brazilian National Council for Scientific and Technological Development–CNPq* (Productivity Fellowship PQ-2 No. 306791/2021-8) and the *Carlos Chagas Filho Foundation for Research Support* of the State of Rio de Janeiro (Young Scholar Fellowship No. 259788 and Grant No. 249104). This study was approved by the *Fluminense Federal University Ethics Committee*, permission No. 29720620.8.0000.5243.

1. Introduction

In the last half-decade, mobile instant messaging services (MIMS) have become a matter of concern for governments, civil society, and academic researchers, due to their opacity and the difficulty of monitoring the circulation of content harmful to democracy, such as mis/disinformation and hate speech, which particularly flood the discussion groups they host (**Rossini**; **Stromer-Galley**; **De-Oliveira**, 2020; **Banaji**; **Bhat**, 2019). There is a recent but vast literature that has sought to discuss these platforms, with an emphasis on specific services, such as *WhatsApp* (**Bursztyn**; **Birnbaum**, 2019), *Telegram* (**Willaert** *et al.*, 2022; **Santos**; **Saldaña**; **Tsyganova**, 2021), and *WeChat* (**Wu**; **Wall**, 2019), among others. And although, among these three examples, Russian and Chinese private messaging services equally pose challenges for their respective contexts, it is *WhatsApp*, due to its enormous popularity, especially in non-Western countries such as Brazil and India, that has boosted public debate around issues such as the spread of fake news (**Resende** *et al*., 2019; **Sacramento**; **Paiva**, 2020), and increased distrust in democratic institutions (**Piaia**; **Alves**, 2020), political radicalization (**Evangelista**; **Bruno**, 2019), and dangerous speech (**Saha** *et al*., 2021; **Matamoros-Fernández**, 2020). In all these cases, there is a lot of discussion about strategies to limit the mass dissemination of certain contents and technical solutions to contain damage to democracy (**Resende** *et al.*, 2019), but little or nothing has been discussed about the effects of environmental opacity on platform cultures, and values shared by users of these services, nor on the practical challenges for implementing democratic controls and monitoring these media.

This article seeks to explore the intrinsic relationship between opacity and transparency and between privacy and publicity, arising from political use and academic research on mobile instant messaging services. Our main goal is to discuss the challenges posed by private messaging services, in particular *WhatsApp*, to the context of democratic transparency, and how to face them. Therefore, the article is based on three different sections.

In the first one, we will present a brief contextualization about how *WhatsApp* has become one of the main protagonists in the Brazilian political scenario (**Moura**; **Michelson**, 2017), and how other countries have also faced situations arising from the way in which dissent groups have made use of the service (**Al-Zidjaly**, 2017). Brazil, together with India, constitutes an exemplary case for analysis. The country has the second largest user-base of the service across the globe, and was perhaps the first to face a major setback due to the spread of fake news and attacks on democracy in *WhatsApp* public discussion groups. In India's parliamentary elections in 2019 (**Garimella**; **Eckles**, 2020), and, in the same year, in Indonesia's general elections (**Baulch**; **Mataromos-Fernández**; **Suwana**, 2022), a similar effect was feared, and Brazil was evoked as a negative example in several circumstances (**Murgia**; **Findlay**; **Schipani**, 2019). Therefore, in this first section of the article, we aim to discuss how and why *WhatsApp* has raised concerns in different democracies.

In the second section, we will put *WhatsApp* and other digital platforms in context, in order to deepen a debate around what the literature has tried to call algorithmic transparency (**Diakopoulos**, 2014). More specifically, we intend to discuss the content regulation and moderation policies assumed by the platform itself and the effects of its actions on the users. Since 2018, in Brazil, *WhatsApp* has incorporated a series of restrictions on the forwarding of messages (**Porter**, 2020), promoted the scale banning of several users (**Mari**, 2019), and has sought to develop institutional partnerships with state agents, such as the *Superior Electoral Court* (*TSE*, 2022). As a counterpoint, one of its main rivals, the Russian *Telegram*, has shown itself to be much more reluctant to participate in this negotiation circuit. The question that remains is: have the efforts carried out by *WhatsApp* really helped to reduce the environmental opacity bequeathed by the service?

Finally, in the third section, we intend to review some of the studies developed on the uses of *WhatsApp* in political contexts. This time, however, instead of focusing on debating how *WhatsApp* has dealt with aspects concerning the privacy of its users, we focus on understanding what challenges this opacity model poses to researchers who deal directly with private data, and often in environments hostile to scholarly research. Thus, if in the previous section we discussed platform transparency, here we discuss what we can call methodological transparency around mobile instant messaging services. In the end, we present some contributions to the theoretical and methodological debate in the fields of political science and political communication regarding the research agenda concerned with MIMS.

2. Private messaging as menace or redemption for liberal-democracies

Private messaging services are nothing new. Applications such as *ICQ*, *AIM* or *MSN Messenger* were extremely popular in the second half of the 1990s. Instant messengers, however, were gradually replaced and incorporated as a functionality of social network sites (SNS), to the point that the uses of this type of platform have shrunk deeply in some countries in the early decades of the 2000s (**Barot**; **Oren**, 2015). Mobile devices, however, have updated the offer for similar services and, fortuitously, new applications have been created in periods of widespread political upheaval in different parts of the globe.

Between 2005 and 2010, the decline in instant messaging tools was partly accompanied by the decline in popularity of some services offered by large news portals, such as *AOL*, which owned *AIM*, and held more than 50% of the private communication market at that time (**Barot**; **Oren**, 2015), and *Yahoo!*, which owned *Yahoo! Messenger*. At the same time, SNSs offered complementary

> This article explores the intrinsic relationship between opacity and transparency and between privacy and publicity, arising from political use and academic research on mobile instant messaging services

private messaging features through the so-called direct messages (DM), still present nowadays on platforms such as *Facebook* (which hosts *Messenger*) and *Twitter*. But the growing popularity of mobile devices introduced a new kind of application whose immediate emphasis was on the quick exchange of messages between users via their respective cell phones. In parallel, the growth in the number of smartphones, the spread of broadband internet and high-speed mobile networks, and the subsequent development of new voice over IP (VoIP) protocols, made instant messaging applications powerful communication tools.

The volume of data exchanged through chat apps surpassed native short-message services (SMSs) for the first time in 2013 (**Barot**; **Oren**, 2015; **Church**; **Oliveira**, 2013), and, in Brazil, this same milestone represents an important change in habits in the population. Data from the *Brazilian Internet Steering Committee* (*https://cgi.br*) regarding online users' activities show that the use of social network sites remains between 70 and 75% among the Brazilian population, from 2011 to 2018, while the use of instant messengers grew from 70 to 92% in the same period (**Chagas**, 2022). The number of people who send instant messages over the internet, according to the survey, is the highest among all other habits, such as sending emails, using blogs, and online forums.

Between 2011 and 2014, several countries experienced mass protests. Symbolic uprisings such as the *Arab Spring*, *Occupy Wall Street*, *Los Indignados* and anti-government demonstrations in Brazil, Chile and Russia, among other examples, have drawn attention to the connection between the use of digital platforms and political participation (**Bennett**; **Segerberg**, 2012; **Klein-Bosquet**, 2012; **Mendonça** *et al*., 2019). Literature has concluded that the use of social networks as news sources, the expression of political opinions, and activism itself, including mobilization through this type of platform, are some of the factors that increase participation through digital media (**Valenzuela**, 2014).

In Brazil, the use of private messaging services has grown exponentially in recent years, in the wake of what has become known in the country as the *2013 June Journeys* (**Chagas**, 2022). Social networking sites and particularly instant messengers have been widely adopted by protesters to organize protests, exchange information about events, and even share memes, as seen in other countries (**Mendonça** *et al*., 2019).

Although there are no reports of state repression and social media regulation actions that justify any degree of mistrust, as there is in the Chinese context (**Mina**, 2019), the adoption of end-to-end encryption systems by MIMS was a crucial component in increasing adoption in this kind of application. Added to this is a no less important economic factor: the expansion of the mobile network in the country, following the privatization of phone companies in the late 1990s. Competition among phone companies in Brazil led to the popularization of prepaid subscription plans among users, who then began to offer discounts or even exemption from expenses for the use of certain services, among them *WhatsApp*. It is worth remembering that this kind of practice is contrary to the provisions of the *Marco Civil da Internet* (*Law No. 12,965/2014*), which establishes equal treatment for all services incorporated by telephone companies. The so-called "zero-rating" plans were, perhaps, the main responsible for the wide penetration of *WhatsApp* as one of the most installed apps on mobile phones across the country (**Evangelista**; **Bruno**, 2019). The result of this is that, today, a wide layer of the popular classes not only use private messaging services as a way of communicating, but, to a large extent, depend on them for work. They are small traders, local markets, delivery people, self-employed professionals and cooperatives who use *WhatsApp* daily as a work tool, and end up exposed to other uses.

According to the company's own data, Brazil has more than 120 million users in 2017, and represents a share of 8% of users worldwide (**Chagas**, 2022). The data is reasonably uncertain, but the percentage of cell phones with *WhatsApp* installed in Brazil is very high. According to *Yahoo! Finance* (2021), 91% of smartphones in the country have *WhatsApp*, which places Brazil in seventh place among the largest users. In Latin America, Argentina (93%) and Colombia (92%) appear ahead. And among the five countries with the largest installed base, three are African, with Kenya ranking first (97%). In Europe, only Turkey and Spain (both with 88%) appear in the top positions.

According to data from *Panorama Mobile Time/Opinion Box Report* (2020), *WhatsApp* is installed on more than 99% of cell phones in Brazil. If one crosses this data with the annual survey released by the *Getulio Vargas Foundation* (**Meirelles**, 2022), according to which Brazil currently has more than one smartphone per inhabitant, and a total of 234 million devices in use, the resulting user base is actually impressive. In terms of number of active users, Brazil is second only to India, which has 390 million accounts (**Iqbal**, 2022).

News consumption on *WhatsApp* is also reported by users as an increasingly relevant activity. In Brazil, more than 50% of the population claimed to use the app as a news source, in 2019 (**Newman** *et al.*, 2019). A year earlier, in 2018, another survey stated that 62% of the Brazilian population believed in the news they received via *WhatsApp*, while only 8% did not (**Passos**, 2018). And yet another last survey, also from 2018, found that, among Bolsonaro voters, six out of ten individuals obtain information mainly through *WhatsApp* and share political news through groups on the app (*Datafolha*, 2018).

> "According to data from *Panorama Mobile Time/Opinion Box Report* (2020), *WhatsApp* is installed on more than 99% of cell phones in Brazil"

The information consumption habits of the Brazilian population have significantly changed recently and digital platforms in general, and private messaging services such as *WhatsApp* in particular, play an important role in this process. **Rossini** *et al*. (2021), for example, draw attention to how *WhatsApp* has become central to how Brazilians access political information and how they engage politically. And while the authors have suggested elsewhere that *WhatsApp* users who share dysfunctional information are more subject to social correction (**Rossini**; **Stromer-Galley**; **De-Oliveira**, 2020), the concern over how *WhatsApp* has become a pervasive media capable of unbalancing not only the informative diet but the democratic environment itself has been evidenced in many different studies. And not just in Brazil.

In India, one of the most notable effects discussed in the literature is the creation of surveillance networks that facilitate and encourage lynchings based on alleged complaints received through viral messages (**Mukherjee**, 2020; **Banaji**; **Bhat**, 2019). Digital vigilantism, of course, is not exclusive to mobile instant messaging services, but the fact that these users are part of a network with high social capillarity has led Indian authorities to determine that the platform should share user metadata for lawful communication and surveillance protocols (**Arun**, 2019). The panorama reflects what **Phillips** and **Milner** (2020) argue about the deep memetic frames used by extremist groups in the hope of fostering moral panic in the population. The authors claim that such frames have always been a moralizing instrument and have often been used by conservative groups to mobilize the population to react. The difference posed by the digital media regime resides in the fact that, unlike what happened before, in which these moral surges were confined to specific regions and neighborhoods, now, say the authors, information is de-quarantined and circulates widely.

In this sense, studies such as those by **Santos** *et al*. (2019) show important circulation patterns in misinformation messages on *WhatsApp*. The authors analyze how messages that evoked an alleged electoral fraud were spread virally and gained scale in geometric progression with the combined use of *WhatsApp* as a private message service and as a broadcast communication tool, from public discussion groups that include, each, up to 256 users.

Vermeer *et al*. (2020) also recall that *WhatsApp* seems to favor the mobilization of users. And **Gil de Zúñiga**, **Ardèvol-Abreu** and **Casero-Ripollés** (2019) suggest that *WhatsApp* has a positive influence on activism and political participation among its users. **Chagas** *et al*. (2022) align with this conclusion by drawing attention to the participatory distortion effects that the platform plays through calls to action for voting in public consultations, once again, broadcasted through political discussion groups.

Although the literature is not consensual and is somehow inconclusive around the *WhatsApp* electoral impact (**Schaefer** *et al*., 2019), it seems clear that the public discussion groups provide an effect of radicalization in users, since they become part of a kind of echo chamber (**Evangelista**; **Bruno**, 2019). Radicalization is a clear result of the mix between homophilic audiences and hyper-partisan informative diets, as pointed out by studies such as those by **Mont'Alverne**, **Mitozo** and **Barbosa** (2019) and **Santos**, **Chagas** and **Marinho** (2022). But none of this would be possible without an environment of extreme opacity.

Arun (2019) argues that the same technological structure that protects users from eventual invasions of privacy also results in an environment that encourages the dissemination of harmful speeches and online rumors. End-to-end encryption would therefore be both the solution and the very source of the problem. The fact is that the lack of transparency regarding the metadata of messages circulating on *WhatsApp* is reflected in an environment of extreme surveillance and virtually no possibility of democratic moderation and regulation. The most evident symptom of this has been the growing use of *WhatsApp* for the performance of influence operations and astroturfing practices (**Chagas**, 2022), in which agents from the professional field of politics covertly act as spontaneously mobilized audiences. Thus, the platform's lack of transparency results in actions and behaviors that are not only inauthentic but misleading. In the next section, we discuss these effects a little further.

3. The dilemma of privacy and public transparency within *WhatsApp*

WhatsApp was created in 2009, months before some of its current main competitors such as *Viber* (2010-), *Line* (2011-), *WeChat* (2011-), *Telegram* (2013-), and *Signal* (2014-). In February 2014, the service was acquired for US$19 billion by *Facebook, Inc.* (currently *Meta Platforms*), in what was then the largest acquisition of a venture-capital-backed company in history. And, in November of the same year, through a partnership with the company *Open Whisper Systems* (*OWS*), *WhatsApp* announced the implementation of an end-to-end encryption system for all its clients, based on the protocol developed by another instant messenger, *Signal*.

According to the end-to-end encryption system, only the sender and the recipient have access to the messages and contents shared. The company claims that not even it has access to the messages, which would prevent *Meta* itself from moderating or censoring the content circulated, and/or tracking habits and messages for target ads, for instance. Although this is a discussion not yet completely resolved, since some studies and reports claim that the company has access to encrypted content shared by users (**Freitas**, 2019), this technological model added to the adopted discourse has allowed the company to avoid possible charges for a more proactive action in cases of circulation of disinformation and hate speech.

There are several kinds of encryption systems. The simplest models are called symmetric ciphers, when a secret key is shared between the sender and the receiver so that the encrypted message is interpreted. A more complete model is called an asymmetric cipher, where receiver and sender have a public key and a private key. In this way the text is en-

crypted in the public key between the two users, but it can only be decrypted with the private key of each one. This late model prevents third parties from accessing the whole message base, in case of interception. That is, if a third party discovers the private key, it will only have access to a single message, not the entire system (**Teixeira**; **Sabo**; **Sabo**, 2017). The encryption model adopted by *WhatsApp*, however, is a bit of a mystery.

> In short, from a corporate perspective, *WhatsApp* oscillates between transparency and privacy inconsistently. For users of their business accounts, privacy is not considered, but for individual users who disseminate messages on a large scale through groups or broadcast lists, the service reserves the right not to act, claiming the right to privacy of its user base

WhatsApp shares public information about its encryption system in a general way, giving transparency only to the protocol model. However, it does not make clear how the system works in practice within the application. Making opaque, for example, the information whether or not the application stores user data. Although *Meta* claims that it does not have access to forwarded messages, there is no reliable credibility that this does not actually happen, taking into account the company's history, in cases such as *Cambridge Analytica*, where *Facebook* user data were shared with third parties for targeting political ads. In addition, a feature recently adopted by *WhatsApp* as a way to prevent and highlight content spread virally, a tag that allows one to identify frequently forwarded messages, suggests that even encrypted messages can be accessed and identified by the company.

Another point of concern is the adoption by *WhatsApp* company, in 2021, of privacy policies subordinated to local contexts and norms (*Abraji*, 2022). The side effect of this change is that, in European Union countries, governed by the *General Data Protection Regulation* (*GDPR*), privacy rules have become excessively strict, but in countries like Brazil, for example, the new policy has made it easier to share user data with the *Facebook* application. Thus, registration information, phone number, IP address, mobile device information such as model, battery level, signal strength, messaging app version, default browser information, mobile network, and even data transactions and payments, and navigational habits (time, frequency and duration of activities, performance reports, etc.) are now interoperated. In addition, the contents of messages exchanged through *WhatsApp* business accounts are no longer encrypted and the metadata generated can be used by *Facebook* to propose ads (**Cosseti**, 2021).

In short, from a corporate perspective, *WhatsApp* oscillates between transparency and privacy inconsistently. For users of their business accounts, privacy is not considered, but for individual users who disseminate messages on a large scale through groups or broadcast lists, the service reserves the right not to act, claiming the right to privacy of its user base.

This inconsistency has been sustained on different occasions by the rhetoric adopted by the *WhatsApp* office in Brazil in public speeches in which it is emphasized that approximately 90% of the messages exchanged through the application are limited to interpersonal communications between two users, and that the average number of users in groups hosted by the service averages around seven individuals (D. Durigan, personal communication, October 27th 2020). And, although these data are replicated in studies that, having received funding from *WhatsApp*'s own public calls, argue that research on the political effects of dysfunctional information sharing should not emphasize large groups dedicated to political discussion, since they would be the exception and not a rule, and would not reflect the experience of most users (**Rossini** *et al*., 2021), it is notable that the problem is not related to the interpersonal exchanges but to the virality of messages shared by and within these groups (**Santos** *et al.*, 2019). Thus, ignoring the threat posed by end-to-end encryption applied to public groups with the participation of agents from the professional field of politics (**Chagas**, 2022; **Chagas**; **Modesto**; **Magalhães**, 2019), is equivalent to color-blind the nuances acquired by the service by incorporating multiple affordances. Furthermore, as studies such as those by **Santos**, **Chagas** and **Marinho** (2022) and others have tried to demonstrate, *WhatsApp* works mainly as a kind of information hub, integrating different platforms and social groups. In this way, even though the large political discussion groups are absolutely minority, they act in the sense of dynamizing and spreading content to smaller groups and individual users, in a very significant way.

WhatsApp can no longer be seen as just a one-to-one instant messenger, in a private context, but as a very complete broadcasting tool. **Vieira** *et al*. (2020) claim that with the expansion of the application, and especially with the massive use of resources such as forwarding messages to multiple users and groups, the messenger stopped being a merely technological platform and became a media platform.

> "As a media platform, just like a radio or television channel, it spreads content through its information broadcast functions and allows messages to become viral" (**Vieira** *et al*., 2020, p. 4).

One of the pernicious effects of this policy adopted by *WhatsApp* for its discussion groups is the opacity regarding operations such as the mass-messaging, spamming, and the spread of fake news and hate speech (**Resende** *et al*., 2019). As argued below, this opacity is even greater due to the fact that the platform does not have an API and has not sought to facilitate scholarly monitoring and research initiatives.

In response to serious criticism after the 2018 presidential elections in Brazil, *WhatsApp* claims that it has regularly and automatically banned or blocked a series of accounts that violate the platform's rules or current electoral legislation. In

parallel, the app has sought to develop partnerships with fact-checking organizations such as *AFP Checamos*, *Agência Lupa*, *Aos Fatos*, and *Estadão Verifica*. In India, similar initiatives involve partnerships with *Digit Eye*, *Fact Crescend*, *Factly*, *India Today*, *Newschecker*, *Newsmobile*, *The Healthy Indian Project*, *The Quint - WebQoof*, and *Vishvas News*. There are also fact-checkers that offer the same kind of service in Albania, Argentina, Colombia, Croatia, Ecuador, France, Germany, Georgia, Ghana, Greece, Guinea, Indonesia, Ireland, Italy, Ivory Coast, Kenya, Mexico, Nigeria, Peru, Portugal, Senegal, Spain, South Africa, Sri Lanka, Turkey, United Kingdom and United States (*WhatsApp*, 2022).

Partnerships have also been carried out with public institutions. In Brazil, the *WhatsApp* office signed a memorandum of agreement with the *Superior Electoral Court* (*TSE*) in which it commits to implement or assist in the implementation of initiatives to combat misinformation about the electoral process. Among the planned actions, *WhatsApp* proposed to:

- hold seminars for *TSE* and *Regional Electoral Courts* (*TRE*s) on the application;
- produce booklets about the service; and
- assist in the implementation of actions for the rapid identification and containment of misinformation, such as the creation of an extrajudicial channel for reporting content that violates legislation and the development, in partnership with the *TSE*, of a chatbot with information about elections from reliable sources (*TSE*, 2022).

However, from a technical point of view, the measures taken by *WhatsApp* so far to contain these effects have focused on limiting forwarding messages and blocking/banning users who disseminate infringing content. Little or nothing has been done to make the environment more transparent, quite the opposite. Among the most recently incorporated affordances, *WhatsApp* allowed:

- blocking messages from unknown numbers and reporting them as spam;
- setting which users can add a given account to new groups.

In addition, since 2018, *WhatsApp* has limited the forwarding of the same message to only 20 contacts (including groups) at a time in Brazil, and, in 2019, it again reduced this limit to five contacts. In 2020, the service started to identify high-frequency messages (HFM) with a specific label, and then limited HFM forwarding from five to just one contact at a time. These limits reportedly reduced message sharing in Brazil by 30%, and reduced high-frequency message traffic worldwide by 70% (*WhatsApp*, 2021). However, there is still little transparency regarding the measures taken after a message has been reported as spam, and there is not, so far, much clarity also about accounts banned from the platform for misbehavior. According to the company's own international reports, an average of 8 million accounts are banned from the platform monthly and about 95% of the deletions are made by automatic detection (**Bento**, 2022).

These measures proved to be ineffective mainly in a context where the dissemination of misinformation has taken on a wide scale and has been ideologically oriented, such as in Brazil. Public political discussion groups are mostly managed by Brazilian far-right supporters (**Chagas**, 2022), and, although sustained by users who show a high degree of engagement, they are also marked by an expectation of horizontality, in which moderation allows for the free sharing of messages, as long as it meets the ideological prerogative of the group itself, that is, only content that is politically in disagreement with the group's guidelines is summarily eliminated.

WhatsApp does not have the same structural features as other social network sites. For example, it does not have public profiles, it does not publicly display connections between users, and it does not have affordances recently incorporated by different digital platforms, such as the timeline that organizes the posts displayed through a social algorithm, according to personal and navigation preferences for each user. According to **Sahafizadeh** and **Ladani** (2020) the use of instant mobile messengers, such as *WhatsApp*, has constituted a new model for online communication. Unlike traditional social networks, based on friendship relations, MIMS users face two different modes of communication:

- peer-to-peer, they usually need to have the contact of their interlocutor registered in their devices;
- in discussion groups, the platform works as a broadcast environment, in which it is not necessary to know the people with whom they interact.

Much of the discussion focused on the danger that social media poses to contemporary democracies emphasizes the opacity surrounding these social algorithms and content recommendation systems. For instance, there are studies that discuss how *YouTube* recommends videos to its users based on non-transparent criteria, which results in recommendations for hyper-partisan or harmful content (**Bryant**, 2020). *Twitter* has already admitted that far-right content has been favored and gained wide exposure through its platform (**Huszár** *et al*., 2021). And *Meta* and *Facebook* have similarly been subjected to public scrutiny on different occasions, including the repercussions of the *Cambridge Analytica* scandal. **Benkler**, **Faris** and **Roberts** (2018) also sought to demonstrate that the affordances of digital platforms have been used to pri-

> Unlike traditional social networks, based on friendship relations, MIMS users face two different modes of communication. In the first, peer-to-peer, they usually need to have the contact of their interlocutor registered in their devices. But, in discussion groups, the platform works as a broadcast environment, in which it is not necessary to know the people with whom they interact

vilege the exposure of content by politically interested groups, and **Woolley** and **Howard** (2018) draw attention to how computational resources, including bots and sock-puppet networks, have been used to spread political propaganda in digital media.

Literature, however, rarely focuses on *WhatsApp*. Not only because of the difficulty of monitoring the platform that it poses for researchers, as we discuss below, but also because its architecture is fundamentally different from that of social media, and, because of that, the same criteria that apply to the demand for greater algorithmic transparency as advocated by various scholars, do not apply to *WhatsApp*'s opacity environment.

In the case of *WhatsApp*, the main problem is not the opacity of its algorithms, but the opacity in the metadata of its contents and users. Briefly, it is impossible to determine how many political discussion groups there are on the platform, or how a particular message circulated, if it is overperforming in engagement, or even if it was viewed above average. It is not possible to identify which links have circulated the most and from which information sources (**Santos**; **Chagas**; **Marinho**, 2022; **Mont'Alverne**; **Mitozo**; **Barbosa**, 2019), nor which memes or anti-science content have been passed on among users and which users are the ones who share such content (**Massuchin** *et al*., 2021). As a result, unlike what happens in other social media, in *WhatsApp*, the main problem is not the influence of opaque systems for recommending content to users, but the complete absence of parameters capable of guiding users about the content that reaches them through direct recommendation from other users, which we are calling here environmental opacity.

We do not know with absolute transparency how the end-to-end encryption adopted by the service operates, and whether or not *WhatsApp* stores messages on its own server. We don't know which accounts are banned and for what reasons. And we don't even have mechanisms to track harmful messages or reports on the decisions made by the platform.

It is not just about charging algorithmic accountability from *WhatsApp*, as is done with other platforms, therefore (**Diakopoulos**, 2014). The prevention of misinformation and dangerous content on *WhatsApp* should not be limited to technical restrictions or blocking and banning accounts based on automated identification of inauthentic behavior, but above all on the need to provide more transparency to the metadata of content and users on the platform. This type of solution does not violate user privacy principles, as it is not necessary to identify users a priori. But it is perfectly possible from a technical point of view to present, for instance, how many times a given message was forwarded, how many users viewed it, at what date and time it was created, and even what were the global reactions to the message. These elements alone may be insufficient to discern whether it is harmful content or not, but they can certainly help in the decision of who accesses a particular piece. These are precisely the same metrics that we have when we access a *YouTube* video or read a *Facebook* post. But they are not available on *WhatsApp* or other private communication services, even though, as we have seen, such platforms merge private messaging with broadcast mode.

The treatment of these two modes of communication based on the same principles, with the prevalence of privacy over public transparency in the case of discussion groups and broadcast lists, has resulted in a first dilemma concerning transparency on *WhatsApp*, according to which the same characteristics that give the platform a high degree of protection for users' privacy contribute to undermining the democratic environment due to an absolute lack of public transparency. Add to this the enormous difficulty of researchers in penetrating these environments, as we discuss further, and we have the formula for a time bomb to discredit democratic institutions.

From a legal standpoint, one of the pioneering laws to bring *WhatsApp* and traditional social network sites on a par, and punish the ones who create or distribute fake news, requiring a quarterly report on each platform's policies and moderation decisions, is the United Arab Emirates *Cybercrime Law* (**Kabha** *et al*., 2019). According to the authors, although it receives criticism for being too rigorous, the legislation has prevented harmful uses of *WhatsApp* and other social media.

In Brazil, two cases are currently under discussion in the *Supreme Court*, the *Direct Action of Unconstitutionality No. 5.527*, and the *Claim of Non-Compliance with a Fundamental Precept No. 403*. In both cases, the judicial decisions of lower courts to determine the national suspension of message exchange services are questioned. The docket reports of both actions, however, present distinct perspectives. In the first case, Justice Edson Fachin opposes to the suspension of apps by court orders. In the second, Justice Rosa Weber understood that end-to-end encryption cannot prevent access to judicial means. Recently also, the *Brazilian Senate* approved the Bill No. 2.630, called the "Fake News Bill", which provided rules for social networks and applications such as *WhatsApp*, to fight disinformation. The project still needs to be approved by the House to be sanctioned.

> "The prevention of misinformation and dangerous content on *WhatsApp* should not be limited to technical restrictions or blocking and banning accounts based on automated identification of inauthentic behavior, but above all on the need to provide more transparency to the metadata of content and users on the platform"

In the Brazilian Executive, the new government of Luiz Inácio Lula da Silva apparently has been moving quickly to contain the circulation of messages of hate and incitement to crimes in the online environment. The now Brazilian Minister of Justice and Public Safety, Flávio Dino, delivered a "Democracy Package" to the *Parliament*. The measures involve

classifying the organization and incitement to anti-democratic demonstrations as a crime of terrorism, such as those that took place in Brazil on January 8 2023, when supporters of former President Jair Messias Bolsonaro invaded and vandalized the headquarters of the three powers of the Republic, in a clear attempt at a coup d'état. The news has drawn attention to the role of MIMS such as *WhatsApp* and *Telegram* in mobilizing for the acts (*Poder360*, 2023).

> "The means to resolve these dilemmas are in the hands of legislators and digital platforms. These agents can ensure that scientific research has minimal conditions to function, and thus provide society with more data and inputs about polarized environments, as is the case with extremist political discussion groups on *WhatsApp*"

On the discussion of regulation, **Medeiros** and **Singh** (2020) argue that lawmakers cannot ignore the negative consequences of encouraging overzealous moderation practices. For forcing changes in platform architecture, specifically the removal of end-to-end encryption, and proactively enforcing content removal responsibilities can be problematic and compromise dissenting discourse. On the other hand, the platforms themselves have relied on this defense to deny the existence of the problem. The point, however, is that conversations between users cannot be confused with permissiveness for anti-democratic mayhems. And this distinction is at the root of the difference between peer-to-peer and broadcast communication in these applications, something that legislation rarely pays attention to.

4. The dilemma of methodological transparency within *WhatsApp*

There are fundamentally two types of difficulties faced by researchers who focus on investigations on *WhatsApp*, which can also be described as methodological transparency issues:

- the first one is the challenge posed to deal with the almost complete lack of transparency on the part of the platform with regard to its data, and
- the second and no less important is the challenge posed by the need to respect the private nature of the data related to the individuals observed.

Regarding the first difficulty, according to **Benevenuto** and **Ortellado** (2020), researchers would benefit if *WhatsApp* regularly publishes two types of information:

- aggregated surveys on platform users, such as the number of users, groups and number of messages distributed in groups, or information about viral messages, such as the number of times a certain content was shared; and
- information and protocols for data collection by academic researchers, i.e. an API documentation.

In recent years, *Twitter* (**Tornes**, 2021), *YouTube* (*YouTube*, 2022) and *TikTok* (**Roth**, 2022) have developed programs aimed at researchers, including specific APIs for the academic audience. Other *Meta* platforms, such as *Facebook* and *Instagram*, have APIs for developers and agreements with some scholarly institutions for the cession of data for academic investigations (**Li** *et al*., 2022).

WhatsApp Business has an API for developers, but the ordinary version of the app, aimed at individual use, lacks greater transparency regarding the procedures for scraping data. Thus, most of the research developed around *WhatsApp* has been anchored either in qualitative analysis strategies, such as ethnography (**Cesarino**, 2020), or in unauthorized data scraping methods (**Piaia**; **Alves**, 2020), which can come up against in the platform's own automated filters aimed at identifying and banishing inauthentic behavior. What often happens is that, as the process involves automation mechanisms, for the data to be collected, the application itself interprets that the action performed is suspicious and inactivates the account that was being used for research, considering that the behavior violates the norms.

An alternative for large-scale data collecting from *WhatsApp* is the app's native chat export tool or the scraping of browsing sessions via the *WhatsApp* Web app. In both cases, researchers face challenges in handling and wrangling the data, as available metadata is scarce –often only the author and content of the message and the date of its publication are available (**Gruber**, 2022).

At the same time, as researchers do not have data on the service's user base and groups, the samples they deal with in their respective investigations are invariably non-probabilistic samples, which makes it difficult to draw inferential conclusions. To circumvent this situation, some researchers have chosen to carry out surveys with individuals who self-identify as *WhatsApp* users, instead of dealing with the data published through the platform and the users associated with it (**Gil de Zúñiga**; **Ardèvol-Abreu**; **Casero-Ripollés**, 2019; **Rossini**; **Stromer-Galley**; **De-Oliveira**, 2020). Something similar happens with viral content. Since the data obtained are not representative in scale of what circulates on *WhatsApp* in general, it is not possible to certify whether a content replicated N-times in a certain set of observed groups actually achieved a significant reach among the entire user base of the service.

Therefore, the first problem related to research on *WhatsApp* and transparency lies in the fact that the platform's environmental opacity compromises the data extraction and collection. On the other hand, the second issue is related, as we stated earlier, to the observed individuals.

This second challenge refers to the private nature of the data that circulates through *WhatsApp*. Unlike what happens on other social media platforms, *WhatsApp* users do not sign a consent form for the publicity of their content. On the contrary, *WhatsApp* bases its experience on a private messaging service. Thus, researchers need to deal with a potential violation of users' privacy when investigating theses environments. Research on *WhatsApp* data, as a rule, has two different expedients in order to avoid compromising privacy rules:

- the first of these is the anonymization of data, which includes the complete de-identification of users, and
- the second is the presentation of results only on an aggregated data scale, thus avoiding individualized analyses.

In the European Union, the *General Data Protection Regulation* (*GDPR*) establishes strict rules regarding the privacy and data protection of citizens residing in member countries. By 2020, Brazil has adopted similar legislation in this regard, called the *General Data Protection Law* (*LGPD*, *Law No. 13,709/2018*), which is a set of regulations for the *Federal Public Administration*, companies and institutions regarding the handling of personal information. The Brazilian *LGPD* establishes as personal data

> "[all] information related to an identified or identifiable natural person,"

and therefore considers sensitive, among others, data related to racial or ethnic identity, gender, religious conviction, political opinion, union affiliation or civil organization, and those relating to health or sex life, genetic or biometric traits of any individual (**Wirnerm**, 2019).

Due to this legal treatment, and also for ethical reasons, research in encrypted private chat apps has sought to ensure full anonymization for research subjects. However, the difficulties do not stop there. Most research ethics councils usually consider the application of an informed consent form to subjects as a good practice. The *WhatsApp* discussion group environment, however, is absolutely volatile, with users coming and going all the time, and discussion groups that are created and suddenly dissolved. Thus, **Barbosa** and **Milan** (2019, p. 59) draw attention to how much this type of platform requires an innovative ethical and methodological approach, in which

> "avoid reducing research ethics to a one-stop checklist [...]; moving past the consent form as the sole and merely regulatory moment of the researcher-research subject relationship."

The authors, however, advocate a research agenda, which

> "embraces transparency and when the research question allows covert methods, avoid dishonest bypasses."

This mention is in line with what **Chagas**, **Modesto** and **Magalhães** (2019) discuss about covert research protocols. In Brazil, covert research is authorized by *Resolution No. 510, of April 7, 2016*, of the *National Health Council* (*CNS*), a collegiate body of the *Ministry of Health* that has an intersectoral commission responsible for implementing norms and guidelines for research involving human beings, the *National Research Ethics Commission* (*Conep*). All research with human subjects, from any area, must be appreciated and evaluated by this body prior to its development. *Resolution No. 510* establishes that covert research is the one

> "conducted without the participants being informed about the objectives and procedures of the study, and without their consent being obtained before or during the research,"

and it is justified only

> "in circumstances in which the information about goals and procedures would change the target behavior of the study or when the use of this method is presented as the only way to conduct the study."

It is a methodological approach reserved for liminal situations, in which individuals cannot even recognize themselves as observed subjects, since this would alter their usual behavior. It is a different procedure, therefore, from that of clinical studies in which substances such as placebo are administered to patients, since, in these cases, consent for the administration of the drug or vaccine is required, even if it is applied only in one portion of the research subjects.

Barbosa and **Milan** (2019) recommend avoiding this type of strategy as much as possible, as well as **Padilha** *et al*. (2005), who argue that this research approach suppresses the right of subjects not to be researched, but they assent that there are scenarios in which data collection in other ways is simply unfeasible. **Chagas**, **Modesto** and **Magalhães** (2019) claim that, especially, research agendas developed in far-right private communication groups require a little more flexibility with precepts of research with human beings. They highlight that in cases like those covert research is often necessary. This is because, in extremist groups, it is common for the simple presentation of the researcher to lead to an immediate expulsion. It is a hostile field for academic research, and absolute methodological transparency is not always able

> "The first dilemma concerns the relationship that such services have sought to establish between the privacy of their users and public transparency, piggybacking on the rhetoric of privacy protection to deny an environment of transparency regarding metadata of potentially harmful content and inauthentic behavior. The second dilemma concerns an effect of the first, weakening scholars due to the private nature of the data available for research"

to resolve the effects of political radicalization. This finding leads to a second important dilemma concerning transparency in *WhatsApp*, according to which more methodological transparency is not always able to generate consent from the research subjects, and environments resistant to science may eventually demand exceptional treatment in this regard.

5. Last remarks

This article aimed to discuss issues involving mobile instant messaging services (MIMS), in particular *WhatsApp*, and transparency. We based our observations on two levels, which we called dilemmas:

The first dilemma concerns the relationship that such services have sought to establish between the privacy of their users and public transparency, piggybacking on the rhetoric of privacy protection to deny an environment of transparency regarding metadata of potentially harmful content and inauthentic behavior.

The second dilemma concerns an effect of the first, weakening scholars due to the private nature of the data available for research. In the latter case, although methodological transparency is a highly desirable requirement, there are situations in which research demands covert methodological approaches, in order to ensure that certain arenas are not completely impenetrable.

These two dilemmas have represented important difficulties for the development of academic research on *WhatsApp* in recent years. Even so, the advances are remarkable, as most of the bibliographic references mobilized throughout this text are able to show.

It should be noted that research ethics, especially in the digital humanities, cannot be reduced to completely inflexible instruments that do not respect specific contexts and situations. On the other hand, it is not our intention to prescribe a playbook without norms, in which the game played equates researchers with anti-science radicals. Instead, what we suggest is that the means to resolve these dilemmas are in the hands of legislators and digital platforms. These agents can ensure that scientific research has minimal conditions to function, and thus provide society with more data and inputs about polarized environments, as is the case with extremist political discussion groups on *WhatsApp*.

Within this context, this article claims that it is absolutely urgent that States develop regulatory models to cope with the spread of disinformation and hate speech in digital platforms. While the major concern of companies has been a compromised and controversial reading of the principle of freedom of expression, little or no transparency about their own actions and business models has been provided to public authorities. MIMS platforms should be called to transparently explain how their encryption system works to users; regularly report to authorities and civil society basic information about banned accounts and the motivations for such decisions; also regularly report the metadata of public groups and frequently forwarded messages; provide users with metadata on the circulation of the viral messages, their senders, and their engagement metrics; and, finally, provide an API for researchers with clear and transparent metadata available.

Other more heterodox actions can also be prompted, but should equally provide guarantees for freedom of expression and protection of personal data. Among them may occur: the enforcement of moderation policies; individual and group sanctions such as alerts, content removal, and user deplatforming. In all cases, however, it is important to bear in mind that it is not a matter of adopting autocratic methods to prevent anti-democratic manifestations. The most important thing, in all scenarios, is to encourage platforms to dialogue with public authorities, civil society, and researchers. Instead of surrounding themselves with an opaque model, perhaps adopting more transparency, it leads to a win-win conclusion.

> Los medios para resolver estos dilemas están en manos de los legisladores y las plataformas digitales. Estos agentes pueden garantizar que la investigación científica tenga unas condiciones mínimas para funcionar, y así proporcionar a la sociedad más datos y aportaciones sobre entornos polarizados, como es el caso de los grupos de discusión política extremista en *WhatsApp*

6. Referencias

Abraji (2022). *O papel das plataformas digitais na proteção da integridade eleitoral em 2022*. Associação Brasileira de Jornalismo Investigativo (Abraji), Book Amazon.
https://goo.su/3stjyP

Al-Zidjaly, Najma (2017). "Memes as reasonably hostile laments: a discourse analysis of political dissent in Oman". *Discourse & society*, v. 28, n. 6, pp. 573-594.
https://doi.org/10.1177/0957926517721083

Arun, Chinmayi (2019). "On *WhatsApp*, rumours, and lynchings". *Economic & political weekly*, v. 54, n. 6.
https://www.epw.in/journal/2019/6/insight/whatsapp-rumours-and-lynchings.html

Banaji, Shakuntala; **Bhat, Ram** (2019). "*WhatsApp* vigilantes: an exploration of citizen reception and circulation of *WhatsApp* misinformation linked to mob violence in India". *Blog Media@LSE*, 11 November.
https://goo.su/XPcQ

Barbosa, Sérgio; **Milan, Stefania** (2019). "Do not harm in private chat apps: ethical issues for research on and with *Whatsapp*". *Westminster papers in communication and culture*, v. 14, n. 1, pp. 49-65.
https://doi.org/10.16997/wpcc.313

Barot, Trushar; **Oren, Eytan** (2015). *Report guide to chat apps*. Columbia Academic Commons, Tow Center for Digital Journalism Publications.
https://towcenter.gitbooks.io/guide-to-chat-apps/content

Baulch, Emma; **Matamoros-Fernández, Ariadna**; **Suwana, Fiona** (2022). "Memetic persuasion and whatsappification in Indonesia's 2019 presidential election". *New media & society*, Online first.
https://doi.org/10.1177/14614448221088274

Benevenuto, Fabricio; **Ortellado, Pablo** (2020). "WhatsApp data that could help research on misinformation in tackling misinformation: what researchers could do with social media data". *Harvard Kennedy school misinformation review*, v. 1, n. 8, pp. 6-7.
https://doi.org/10.37016/mr-2020-49

Benkler, Yochai; **Faris, Robert**; **Roberts, Hal** (2018). *Network propaganda: manipulation, disinformation, and radicalization in American politics*. Oxford, UK: Oxford University Press. ISBN: 978 0 190923624

Bennett, W. Lance; **Segerberg, Alexandra** (2012). "The logic of connective action". *Information, communication & society*, v. 15, n. 5, pp. 739-768.
https://doi.org/10.1080/1369118x.2012.670661

Bento, Gabrielly (2022). "WhatsApp baniu 2,4 milhões de contas na Índia em julho". *Olhar digital*, 5 setembro.
https://olhardigital.com.br/2022/09/05/seguranca/whatsapp-bane-24-milhoes-de-contas-na-india-em-julho

Bryant, Lauren-Valentino (2020). "The *YouTube* algorithm and the alt-right filter bubble". *Open information science*, v. 4, n. 1, pp. 85-90.
https://doi.org/10.1515/opis-2020-0007

Bursztyn, Victor S.; **Birnbaum, Larry** (2019) "Thousands of small, constant rallies: a large-scale analysis of partisan *Whatsapp* groups". In: *Proceedings IEEE/ACM International conference on advances in social networks analysis and mining 2019*, pp. 484-488.
https://doi.org/10.1145/3341161.3342905

Cesarino, Leticia (2020). "Como vencer uma eleição sem sair de casa: a ascensão do populismo digital no Brasil". *Internet & sociedade*, v. 1, n. 1, pp. 92-120.
https://revista.internetlab.org.br/serifcomo-vencer-uma-eleicao-sem-sair-de-casa-serif-a-ascensao-do-populismo-digital-no-brasil

Chagas, Viktor (2022). "*WhatsApp* and digital astroturfing: a social network analysis of Brazilian political discussion groups of Bolsonaro's supporters". *International journal of communication*, v. 16, pp. 2431-2455.
https://ijoc.org/index.php/ijoc/article/view/17296

Chagas, Viktor; **Mitozo, Isabele**; **Barros, Samuel**; **Santos, João-Guilherme**; **Azevedo, Dilvan** (2022). "The 'new age' of political participation? *WhatsApp* and call to action on the Brazilian senate's consultations on the e-cidadania portal". *Journal of information technology & politics*, v. 19, n. 3, pp. 253-268.
https://doi.org/10.1080/19331681.2021.1962779

Chagas, Viktor; **Modesto, Michelle**; **Magalhães, Dandara** (2019). "O Brasil vai virar Venezuela: medo, memes e enquadramentos emocionais no WhatsApp pró-Bolsonaro". *Esferas*, v. 14.
https://doi.org/10.31501/esf.v0i14.10374

Church, Karen; **Oliveira, Rodrigo** (2013). "What's up with *WhatsApp*? Comparing mobile instant messaging behaviors with traditional SMS". In: *Proceedings of mobile HCI 2013 - Collaboration and communication*, pp. 353-371.
https://www.ic.unicamp.br/~oliveira/doc/MHCI2013_Whats-up-with-whatsapp.pdf

Cosseti, Melissa-Cruz (2021). "O *WhatsApp* compartilha dados com o *Facebook*?". *Tecnoblog*, 8 janeiro.
https://tecnoblog.net/responde/o-whatsapp-compartilha-dados-com-o-facebook

Datafolha (2018). "Datafolha: 6 em cada 10 eleitores de Bolsonaro se informam pelo *WhatsApp*". *Veja*, 3 outubro.
https://veja.abril.com.br/politica/datafolha-eleitor-de-bolsonaro-e-o-que-mais-se-informa-por-redes-sociais

Diakopoulos, Nicholas (2014). *Algorithmic accountability: on the investigation of black boxes*. Tow Center for Digital Journalism, 3 December.
http://www.cjr.org/tow_center_reports/algorithmic_accountability_on_the_investigation_of_black_boxes.php

Evangelista, Rafael; **Bruno, Fernanda** (2019). "*WhatsApp* and political instability in Brazil: targeted messages and political radicalisation". *Internet policy review*, v. 8, n. 4.
https://doi.org/10.14763/2019.4.1434

Freitas, Miguel (2019). *WhatsApp nas eleições de 2018: o embate entre a lei, a tecnologia e o direito à privacidade*. Senado Federal.
http://legis.senado.leg.br/sdleg-getter/documento/download/bf52a4a0-ff2b-4f50-b9f8-af9b85c2a099

Garimella, Kiram; **Eckles, Dean** (2020). "Images and misinformation in political groups: evidence from WhatsApp in India". *Harvard Kennedy school misinformation review*, v. 1, n. 5.
https://doi.org/10.37016/mr-2020-030

Gil de Zúñiga, Homero; **Ardèvol-Abreu, Alberto**; **Casero-Ripollés, Andreu** (2019). "*WhatsApp* political discussion, conventional participation and activism: exploring direct, indirect and generational effects". *Information, communication & society*, v. 24, n. 2, pp. 201-218.
https://doi.org/10.1080/1369118x.2019.1642933

Gruber, Johannes B. (2022). "An R package for working with *WhatsApp* data". *GitHub*, 4 October.
https://github.com/JBGruber/rwhatsapp

Huszár, Ferenc; **Ktena, Sofia-Ira**; **O'Brien, Conor**; **Belli, Luca**; **Schlaikjer, Andrew**; **Hardt, Moritz** (2021). "Algorithmic amplification of politics on *Twitter*". *Proceedings of the National Academy of Sciences*, v. 119, n. 1.
https://doi.org/10.1073/pnas.2025334119

Iqbal, Mansoor (2022). "*WhatsApp* revenue and usage statistics 2022". *Business Fapps*.
https://www.businessofapps.com/data/whatsapp-statistics

Kabha, Robin; **Kamel, Ahmad**; **Elbahi, Moataz**; **Narula, Sumit** (2019). "Comparison study between the UAE, the UK, and India in dealing with *WhatsApp* fake news". *Journal of content, community & communication*, v. 10, n. 5, pp. 176-186.
https://doi.org/10.31620/JCCC.12.19/18

Klein-Bosquet, Oliver (2012). "El Movimiento de los Indignados: desde España a Estados Unidos". *El cotidiano*, v. 173, pp. 89-98.
https://www.redalyc.org/pdf/325/32523131010.pdf

Li, Da; **Pyke, Robert**; **Jiang, Runchao**; **Jagadeesh, Kiran** (2022). "Introducing the researcher platform: empowering independent research analyzing large-scale data from *Meta*". *Meta Research Blog*, 11 January.
https://research.facebook.com/blog/2022/1/introducing-the-researcher-platform-empowering-independent-research-analyzing-large-scale-data-from-meta

Mari, Angelica (2019). "*WhatsApp* banned nearly half a million accounts during Brazilian elections". *Zdnet*, November 20.
https://www.zdnet.com/article/whatsapp-banned-nearly-half-a-million-accounts-during-brazilian-elections

Massuchin, Michele; **Tavares, Camila**; **Mitozo, Isabele**; **Chagas, Viktor** (2021). "A estrutura argumentativa do descrédito na ciencia. Uma análise de mensagens de grupos bolsonaristas de *Whatsapp* na pandemia da Covid-19". *Fronteiras - estudos midiáticos*, v. 23, n. 2, pp. 160-174.
https://doi.org/10.4013/fem.2021.232.11

Matamoros-Fernández, Ariadna (2020). "'El negro de *WhatsApp*' meme, digital blackface, and racism on social media". *First Monday*, v. 25, n. 1.
https://doi.org/10.5210/fm.v25i12.10420

Medeiros, Ben; **Singh, Pawan** (2020). "Addressing misinformation on WhatsApp in India through intermediary liability policy, platform design modification, and media literacy". *Journal of information policy*, v. 10, pp. 276-298.
https://doi.org/10.5325/jinfopoli.10.2020.0276

Meirelles, Fernando S. (2022). *Pesquisa do uso da TI - tecnologia de informação nas empresas*. Fundação Getulio Vargas.
https://eaesp.fgv.br/sites/eaesp.fgv.br/files/u68/fgvcia_pes_ti_2022_-_relatorio.pdf

Mendonça, Ricardo-Fabrino; **Ercan, Selen A.**; **Ozguc, Umut**; **Reis, Stephanie-Lorraine-Gomes**; **Simões, Paula-Guimarães** (2019). "Protests as events: the symbolic struggles in 2013 demonstrations in Turkey and Brazil". *Revista de sociologia e política*, v. 27, n. 69.
https://doi.org/10.1590/1678987319276901

Mina, An-Xiao (2019). *Memes to movements: how the world's most viral media is changing social protest and power*. New York, NY: Penguin Random House. ISBN: 978 0 807056585

Mont'Alverne, Camila; **Mitozo, Isabele**; **Barbosa, Henrique** (2019). "*WhatsApp* e eleições: quais as características das informações disseminadas". *Le monde diplomatique Brasil*, 7 maio.
https://diplomatique.org.br/whatsapp-e-eleicoes-informacoes-disseminadas

Moura, Mauricio; **Michelson, Melissa R.** (2017). "*WhatsApp* in Brazil: mobilising voters through door-to-door and personal messages". *Internet policy review*, v. 6, n. 4.
https://doi.org/10.14763/2017.4.775

Mukherjee, Rahul (2020). "Mobile witnessing on *WhatsApp*: Vigilante virality and the anatomy of mob lynching". *South Asian popular culture*, v. 18, n. 1, pp. 79-101.
https://doi.org/10.1080/14746689.2020.1736810

Murgia, Madhumita; **Findlay, Stephanie**; **Schipani, Andres** (2019). "India: the *WhatsApp* election". *Financial Times*, 4 May.
*https://www.ft.com/content/9fe88fba-6c0d-11e9-a9a5-351eeaef6d84_*

Newman, Nic; **Fletcher, Richard**; **Kalogeropoulos, Antonis**; **Nielsen, Rasmus-Kleis** (2019). *Digital news report*. Reuters Institute; University of Oxford.
https://www.digitalnewsreport.org/survey/2019/overview-key-findings-2019

Padilha, Maria-Itayara-Coelho; **Ramos, Flávia-Regina-Souza**; **Borenstein, Miriam-Susskind**; **Martin, Cleusa-Rios** (2005). "A responsabilidade do pesquisador ou sobre o que dizemos acerca da ética em pesquisa". *Texto & contexto - enfermagem*, v. 14, n. 1, pp. 96-105.
https://doi.org/10.1590/s0104-07072005000100013

Panorama Mobile Time/Opinion Box Report (2020). *Mensageria no Brasil*. Mobile Time.
https://www.mobiletime.com.br/pesquisas/mensageria-no-brasil-fevereiro-de-2020

Passos, Paulo (2018). "Metade dos usuários acredita em noticias compartilhadas no *WhatsApp*". *Folha de São Paulo*, 26 octubre.
https://www1.folha.uol.com.br/poder/2018/10/metade-acredita-em-noticias-compartilhadas-no-whatsapp.shtml

Phillips, Whitney; **Milner, Ryan M.** (2020). *You are here: a field guide for navigating polarized speech, conspiracy theories, and our polluted media landscape*. Cambridge, MA: MIT Press. ISBN: 978 0 262539913.
https://direct.mit.edu/books/book/5041/You-Are-HereA-Field-Guide-for-Navigating-Polarized

Piaia, Victor; **Alves, Marcelo** (2020). "Abrindo a caixa preta: análise exploratória da rede bolsonarista no *WhatsApp*". *Intercom: revista brasileira de ciências da comunicação*, v. 43 n. 3, pp. 135-154.
https://doi.org/10.1590/1809-5844202037

Poder360 (2023). "Governo recebeu mais de 100 mil e-mails pelo 8 de Janeiro". *Poder360*, 6 fevereiro.
https://www.poder360.com.br/justica/governo-recebeu-mais-de-100-mil-e-mails-pelo-8-de-janeiro

Porter, Jon (2020). "*WhatsApp* says its forwarding limits have cut the spread of viral messages by 70 percent". *The verge*, 27 April.
https://www.theverge.com/2020/4/27/21238082/whatsapp-forward-message-limits-viral-misinformation-decline

Resende, Gustavo; **Melo, Philipe**; **Souza, Hugo**; **Messias, Johnnatan**; **Vasconcelos, Marisa**; **Almeida, Jussara M.**; **Benevenuto, Fabrício** (2019). "(Mis)Information dissemination in *WhatsApp*: gathering, analyzing and countermeasures". In: *Proceedings of the WWW'19 conference*, pp. 818-828.
https://homepages.dcc.ufmg.br/~fabricio/download/resende-www2019.pdf

Rossini, Patricia; **Baptista Érica-Anita**; **De-Oliveira, Vanessa-Veiga**; **Stromer-Galley, Jennifer** (2021). "Digital media landscape in Brazil: political (mis)information and participation on *Facebook* and *WhatsApp*". *Journal of quantitative description: Digital media*, v. 1.
https://doi.org/10.51685/jqd.2021.015

Rossini, Patricia; **Stromer-Galley, Jennifer**; **De-Oliveira, Vanessa-Veiga** (2020). "Dysfunctional information sharing on *WhatsApp* and *Facebook*: the role of political talk, cross-cutting exposure and social corrections". *New media & society*, v. 23, n. 8. pp. 2430-2451.
https://doi.org/10.1177/1461444820928059

Roth, Emma (2022). "*TikTok* to provide researchers with more transparency as damaging reports mount". *The Verge*, 27 September.
https://www.theverge.com/2022/7/27/23280406/tiktok-researchers-api-transparency-damaging-reports-china

Sacramento, Igor; **Paiva, Raquel** (2020). "Fake news, *WhatsApp* e a vacinação contra febre amarela no Brasil". *Matrizes*, v. 14, n. 1, pp. 79-106.
https://doi.org/10.11606/issn.1982-8160.v14i1p79-106

Saha, Punyajoy; **Mathew, Binny**; **Garimella**; **Kiran**; **Mukherjee, Animeshe** (2021). "'Short is the road that leads from fear to hate': Fear speech in Indian *WhatsApp* groups". In: *Proceedings of the Web conference 2021*, n. 4, pp. 1110-1121.
https://doi.org/10.1145/3442381.3450137

Sahafizadeh, Ebrahim; **Ladani, Behrouz** (2020). "A model for social communication network in mobile instant messagings". *IEEE transactions on computational social systems*, v. 7, n. 1, pp. 68-83.
https://doi.org/10.1109/TCSS.2019.2958968

Santos, João-Guilherme; **Freitas, Miguel**; **Aldé, Alesandra**; **Santos, Karina**; **Cunhna, Vanessa-Cristine-Cardozo** (2019). "*WhatsApp*, política mobile e desinformação: a hidra nas eleições presidenciais de 2018". *Comunicação & sociedade*. v. 41, n. 2.
https://doi.org/10.15603/2175-7755/cs.v41n2p307-334

Santos, Marcelo; **Saldaña, Magdalena**; **Tsyganova, Kesnia** (2021). "Subversive affordances as a form of digital transnational activism: The case of *Telegram*'s native proxy". *New media & society*, Online first.
*https://doi.org/10.1177/14614448211054830_*

Santos, Nina; **Chagas, Viktor**; **Marinho, Juliana** (2022). "De onde vem a informação que circula em grupos bolsonaristas no *WhatsApp*". *Intexto*, n. 53.
https://doi.org/10.19132/1807-8583202253.123603

Schaefer, Bruno-Marques; **Barbosa, Tiago-Alexandre-Leme**; **Epitácio, Sara-de-Sousa-Fernandes**; **Resende, Roberta-Carnelos** (2019). "Qual o impacto do *Whatsapp* em eleições? Uma revisão sistemática (2010-2019)". *Revista debates*, v. 13, n. 3, p. 58-88.
https://doi.org/10.22456/1982-5269.96255

Teixeira, Tarcisio; **Sabo, Paulo-Henrique**; **Sabo, Isabela-Cristina** (2017). "*WhatsApp* e a criptografia ponto-a-ponto: tendência jurídica e conflito privacidade vs. interesse público". *Revista da Faculdade de Direito da UFMG*, v. 71, p. 607-638.
https://doi.org/10.12818/p.0304-2340.2017v71p607

Tornes, Adam (2021). "Product news enabling the future of academic research with the *Twitter* API". *Twitter*, January 26.
https://developer.twitter.com/en/blog/product-news/2021/enabling-the-future-of-academic-research-with-the-twitter-api

TSE (2022). "*TSE* e *WhatsApp* celebram acordo para combate à desinformação nas eleições 2022". *TSE*, Fevereiro 15.
https://www.tse.jus.br/comunicacao/noticias/2022/Fevereiro/tse-e-whatsapp-celebram-acordo-para-combate-a-desinformacao-nas-eleicoes-2022

Valenzuela, Sebastián (2014). "Analisando o uso de redes sociais para o comportamento de protesto: o papel da informação, da expressão de opiniões e do ativismo". *Compolítica*, v. 4, n. 1, p. 13-52.
https://doi.org/10.21878/compolitica.2014.4.1.56

Vermeer, Susan A. M.; **Kruikemeier, Sanne**; **Trilling, Damian**; **De-Vreese, Class H.** (2020). "WhatsApp with politics?!: examining the effects of interpersonal political discussion in instant messaging apps". *The international journal of press/politics*, v. 26, n. 2, pp. 410-437.
https://doi.org/10.1177/1940161220925020

Vieira, Carolina-Coimbra; **Melo, Philippe-de-Freitas**; **De-Melo, Pedro O. S. Vaz**; **Benevenuto, Fabrício** (2019). "O paradoxo da viralização de informação criptografada no *WhatsApp*". In: *Anais do XXXVII Simpósio brasileiro de redes de computadores e sistemas distribuídos*, v. 37, pp. 403-416.
https://sol.sbc.org.br/index.php/sbrc/article/view/7375

Willaert, Tom; **Peeters, Stijn**; **Seijbel, Jasmin**; **Van-Raemdonck, Nathali** (2022). "Disinformation networks: a quali-quantitative investigation of antagonistic Dutch-speaking *Telegram* channels". *First Monday*, v. 27, n. 9.
https://doi.org/10.5210/fm.v27i5.12533

Wimmer, Mirian (2019). "Cidadania, tecnologia e governo digital: Proteção de dados pessoais no estado movido a dados". En: *Comitê gestor da internet no Brasil*. *TIC governo eletrônico: pesquisa sobre o uso das tecnologias de informação e comunicação no setor público brasileiro*. *2019. ICT electronic government*, pp. 27-36.
https://cetic.br/media/docs/publicacoes/2/20200707094309/tic_governo_eletronico_2019_livro_eletronico.pdf

WhatsApp (2021). "Política de privacidade do WhatsApp". *WhatsApp*, Janeiro 4.
https://www.whatsapp.com/legal/privacy-policy/?locale=pt_BR

WhatsApp (2022). "Organizações da *Aliança Internacional de Checagem de Fatos* (*IFCN*) no *WhatsApp*". *WhatsApp*.
https://faq.whatsapp.com/528263691226435/?helpref=uf_share

Woolley, Samuel C.; **Howard, Philip N.** (eds.) (2018). *Computational propaganda: Political parties, politicians, and political manipulation on social media*. Oxford, UK: Oxford University Press. ISBN: 978 0 190931414

Wu, Yan; **Wall, Matthew** (2019). "The ties that bind: How the dominance of *WeChat* combines with *guanxi* to inhibit and constrain China's contentious politics". *New media & society*, v. 21, n. 8, pp. 1714-1733.
https://doi.org/10.1177/1461444819830072

Yahoo! Finance (2021). "Em quais países o *WhatsApp* é mais popular?". *Yahoo Finance BR*, October 18.
https://br.financas.yahoo.com/video/em-quais-pa%C3%ADses-o-whatsapp-130924324.html

YouTube (2022). "YouTube researcher program". *YouTube*.
https://research.youtube