[DRAFT] Considerations for Closed Messaging Research in Democratic Contexts
September 2020, version 1.4

Connie Moon Sehat with Aleksei Kaminski[1]

Executive Summary

"Closed" messaging apps such as WhatsApp and WeChat have grown in reach and adoption in recent years and transformed elections-related communications. However, understanding their impact upon public discussion poses a conundrum for researchers. On the one hand, widely accessible conversations of public importance exist in these spaces, but on the other, the encrypted, private nature of shared messages make the professional ethics of access and collection less than clear. What considerations should be taken into account when conducting research into closed messaging spaces within democratic contexts, in which the individual right to privacy also prevails?[2]

In order to better understand the challenge, this paper presents a review of research practices taking place within these applications with a focus on election-related themes. Two investigating groups are of interest: *public-interest organizations* focused on human rights, democratic elections, and fact-checking that seek real-time or more immediate impact based upon their findings, and *academics* who want to provide longer-term, systematic analysis of the larger political dynamics created by these closed conversations. Sometimes these groups work in

---

[2] While messenger app usage in countries like China is quite dominant, this paper focuses instead on country contexts with multiple party systems and liberal democratic frameworks that have different requirements for public versus private conversation.

collaboration with each other as well as with other types of researchers, such as those based at technology companies.[3]

Between these groups, there are at least four models of research practice, which are summarized below. First, we share background about the reach of closed messaging applications and the motivation for these research efforts. Then, summaries of the four models follow. Based on the models, the final section of the paper offers a set of key questions to help clarify the ethical tensions posed by investigations into closed message discussions.

Following a workshop to explore best practices into closed messaging spaces at MisinfoCon DC in 2018,[4] the initial framework for this paper came from a MisinfoCon Elections workshop in London 2019. Calls for suggestions of related case studies, as well as for researchers who might be interested in collaborating, were issued during the first half of 2020. This paper has so far received limited external feedback as well as clarifications from some investigators cited. The draft will be finalized after further consultation during a workshop on the topic later this year.

Background: The Reach of Closed Messaging Systems

Closed messaging applications or "Messenger Apps" such as WhatsApp, Facebook Messenger, and WeChat are now among the world's leading social media platforms; according to marketing analytics reports, these applications were among the top 10 most-used social platforms at the onset of 2020.[5] Different from the openness of Twitter or even the default Facebook experience, these "private chat" spaces for personal exchange are having a larger impact upon public discourse than originally envisioned.[6]

There are a few reasons for this broad effect: closed messaging chats —which range in size from one-on-one discussion to groups of over 200 or even 100,000 accounts — can be much larger than real-time in-person gatherings. At the same time, the chats are equipped with both the speed of instant messaging and the advantage of asynchronous reach (i.e., participants can still catchup on messages sent hours before and provide new life to a discussion).

There are also a number of reasons for the increase in private messaging usage. For one, these applications enable a measure of encrypted privacy for those who want it — even at the larger scales of group discussion — which is not part of larger open platforms such as Reddit. For those with prepaid data accounts, closed messaging applications also offer a texting, phone, and video service that can work across different countries that either does not require additional cost or is

---

[3] While this topic has relevance for corporate research and development approaches, this paper is focused upon public, outward-facing research results. Two examples of collaboration with corporations are included in the review.

[4] Connie Moon Sehat, "Closed Messaging Spaces: Thinking About Best Practices and Ethics in a Global Environment," *Medium* (blog), October 4, 2018, https://misinfocon.com/closed-messaging-spaces-thinking-about-best-practices-and-ethics-in-a-global-environment-10fc82191743.

[5] For example, Simon Kemp, "Digital 2020: Global Digital Overview" (DataReportal in collaboration with HootSuite and WeAreSocial, January 30, 2020), 95, https://datareportal.com/reports/digital-2020-global-digital-overview.

[6] While closed groups in Facebook offer a similar experience to some of the features of encrypted message chat services, and may warrant similar research approaches, we are focusing on the distinguishing characteristics of applications like WhatsApp and Telegram in this paper.

cheaper than an international plan.[7] Whatever the exact reason, the rise of encrypted direct messaging applications is notable; for the last several years, they are the only type of mobile application in which first-time and continued use has grown. In 2016, for example, *The Economist* noted that teenagers spent more time sending instant messages than they did "perusing social networks."[8]

For a sense of proportion when it comes to closed message application usage: at the time of *The Economist's* reporting in 2016, 1 billion account holders actively utilized Facebook's WhatsApp on a monthly basis. By January 2020, WhatsApp gained an additional 600 million monthly active user accounts or unique monthly visitors, totaling 1.6 billion active users per month; in addition, the application reached 2 billion overall registered accounts in February.[9] Facebook Messenger usage was also very high, totaling 1.3 billion users per month, with WeChat following at 1.1 billion.[10]

These numbers indicate that a significant percentage of Internet users overall are using closed messaging services. In 2018, approximately 49% of the world's population —3.7 billion Internet users — were reported to use the Internet, a number that rose to 4.1 billion or 53.6 percent of the global population by the end of 2019.[11] Therefore, closed messaging applications currently include at least 50% of global Internet users, if not more. Isolation and social distancing resulting from COVID-19 may be contributing to a further spike in global usage: earlier this year, WhatsApp usage appeared to have increased overall by 51%, with Spain reporting a 76% increase in WhatsApp usage and China's WeChat and Weibo reporting increases of 58%.[12]

The popularity of closed messaging application usage has been very popular outside of the US, Australia and certain EU countries; regions with dominant closed message usage include South

---

[7] Examples of reasoning can be found in Jordan McMahon, "Why We Should All Ditch Other Messaging Apps for Signal," *Wired*, November 5, 2017, https://www.wired.com/story/ditch-all-those-other-messaging-apps-heres-why-you-should-use-signal/; Thorin Klosowski, "What's the Deal with All These Messaging Apps?," *Lifehacker* (blog), April 10, 2014, https://lifehacker.com/whats-the-deal-with-all-these-messaging-apps-1561543034; Meghan Keaney Anderson, "Why We're Thinking About Messaging Apps All Wrong," *HubSpot* (blog), February 9, 2016, https://blog.hubspot.com/marketing/messaging-apps.

[8] "Mobile Services - Bots, the next Frontier," *The Economist*, April 9, 2016, sec. Business, https://www.economist.com/business/2016/04/09/bots-the-next-frontier.

[9] Kemp, "Digital 2020: Global Digital Overview"; WhatsApp Blog, "Two Billion Users -- Connecting the World Privately," *WhatsApp Blog* (blog), February 12, 2020, https://blog.whatsapp.com/two-billion-users-connecting-the-world-privately.

[10] Kemp, "Digital 2020: Global Digital Overview."

[11] Susan Halford et al., "Understanding the Production and Circulation of Social Media Data: Towards Methodological Principles and Praxis," *New Media & Society* 20, no. 9 (September 2018): 3341–58, https://doi.org/10.1177/1461444817748953; International Telecommunication Union, "Statistics," *International Telecommunication Union* (blog), accessed June 30, 2020, https://www.itu.int/en/ITU-D/Statistics/Pages/stat/default.aspx.

[12] Pranav Hegde, "Coronavirus Impact | WhatsApp Global Usage Spikes by 51% during Pandemic," *MSN News: Money Control* (blog), March 26, 2020, https://www.msn.com/en-in/news/other/coronavirus-impact-whatsapp-global-usage-spikes-by-51-25-during-pandemic/ar-BB11JFMt; Nic Newman, "Reuters Institute Digital News Report 2020" (Reuters Institute for the Study of Journalism, 2020), https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2020-06/DNR_2020_FINAL.pdf.

America, Africa, and much of Asia.[13] Despite WhatsApp and WeChat being the most widely used Messenger Apps worldwide, other competitors have strongholds in regional markets. Examples include Line and KakaoTalk, instant messaging applications from South Korea, or Rakuten-owned Viber. For example, in 2018, 99.2% percent of South Koreans surveyed by the Korean Ministry of Science and Information and Communication Technology were using KakaoTalk.[14] Line is as of this year the most popular messaging app in Japan and Taiwan, among others.[15]

Given the reach of closed messaging applications, misinformation or false information has definitely created concern. We already know that people are encountering news within these applications with increasing rates: as of 2018 and among people surveyed in 37 countries, the average usage for news in these spaces had more than doubled to 16% in four years.[16] The Reuters Institute for the Study of Journalism at Oxford also reported this year that more than half of respondents (56%) remain concerned about what is real and fake on the Internet when it comes to news circulated online. In countries such as Kenya and Brazil, where closed messaging apps are dominant, respondents reported correspondingly the most concern about misinformation within these particular social media spaces. The same level of concern exists in Chile, Mexico, Malaysia, and Singapore.[17] While the global prevalence of messenger applications has created a central vehicle for communication among billions of individuals, it has also made flow of false information a key worry as well.

Motivations for understanding the political nature of closed messaging spaces

Given their prevalence, it is not surprising that closed messaging applications now factor in election environments. Since 2015, scholars have noted that WhatsApp in particular "has established itself as a powerful political tool for spreading political information," whether misinformation or not.[18] The earlier 100-group limit within WhatsApp during the 2015 elections for the Israeli Knesset (Parliament), for example, seemed to open possibilities for *closer* and informal connections with candidates – a potential advantage for political discussion; similar expressions of closeness between political candidates and constituencies were reported in Indonesia as well.[19] After 2016, especially following the revelations of Internet Research

---

[13] Considering both the top messenger applications in addition to overall global usage, see both We Are Social, "2018 Digital Yearbook," https://www.slideshare.net/wearesocial/2018-digital-yearbook-86862930?ref=https://wearesocial.com/; Kemp, "Digital 2020: Global Digital Overview."

[14] Korean Ministry of Science and ICT, "2018년 인터넷이용실태조사 결과 발표 (Announcement of the 2018 Internet Usage Survey)," *대한민국 정책포털 KOREA.KR 정책브리핑 (Korea Policy Portal KOREA.KR Policy Briefing)* (blog), February 25, 2019, http://www.korea.kr/news/pressReleaseView.do?newsId=156318833.

[15] Kemp, "Digital 2020: Global Digital Overview," 95.

[16] Nic Newman et al., "Reuters Institute Digital News Report 2018" (Reuters Institute for the Study of Journalism, 2018).

[17] Newman, "Reuters Institute Digital News Report 2020."

[18] Sérgio Barbosa and Stefania Milan, "Do Not Harm in Private Chat Apps: Ethical Issues for Research on and with WhatsApp," *Westminster Papers in Communication and Culture* 14, no. 1 (August 14, 2019): 49–65, https://doi.org/10.16997/wpcc.313.

[19] Sharon Haleva-Amir, "Not All about That Facebook: Political Campaigns and Civic Engagement in the 2015 Elections," *Israel Affairs* 22, no. 3/4 (July 2016): 711–26, https://doi.org/10.1080/13537121.2016.1174375; Suwandi Sumartias Sumartono, "Utilization of Whatsapp As A Political Communication Channel Politicians Padang

Agency- and Cambridge Analytica-powered mass manipulation on Facebook, concerns about disinformation within related social media spaces such as WhatsApp and Facebook Messenger seemed warranted as well.[20]

More recently, the pivotal role of WhatsApp in the 2019 Indian legislative elections has exemplified the thirst for leveraging these special spaces for larger political purposes. Users for example were being added to closed groups, sometimes unsolicited and without their consent, until April 2019.[21] New tactics also included the distribution of misinformation and jokes through local-level WhatsApp Bharatiya Janata Party initiatives. Despite the continuation of in-person rallies, television and newspaper coverage, and traditional canvassing, WhatsApp became the country's default mode of communication and preferred vehicle for campaign messaging.[22] As Ankit Lal, a top strategist for the Aam Aadmi Party, opined: "We wrestle on Twitter. The battle is on Facebook. The war is on WhatsApp."[23]

Yet given the closed nature of this space, it is technologically difficult – intentionally so – to explore the nature of these political communications and the extent to which they have a discernable effect upon election outcomes. At the same time, it is unsurprising that journalists, academics, human rights workers, election observers, and other researchers would want to analyze these communication environments better. Reasons include acertaining the needs and desires of citizens or, more immediately, protecting their interests in the face of potential manipulation via disinformation campaigns. Questions from First Draft for journalists, for instance, make clear the range of interests that news stories might have for such communities (see Appendix A).

Cases from the 2018 and 2019 elections in Brazil and India clarify some specific concerns of researchers. First, we have learned that messages from politically oriented groups can be quite slanted in these spaces, and that closed messaging apps take advantage of linkages to other social media platforms. For example, political scientists Camila Mont'Alverne and Isabele Mitozo analyzed a sample of 213 public Brazilian WhatsApp groups with messages containing links to YouTube.[24] Out of these groups, Mont'Alverne and Mitozo found that the largest amount of

---

City," in *Advances in Social Science, Education and Humanities Research*, vol. 136 (2nd International Conference on Social and Political Development, Atlantis Press, 2017), 428–31, https://doi.org/10.2991/icosop-17.2018.67.

[20] Newman et al., "Reuters Institute Digital News Report 2018"; Newman, "Reuters Institute Digital News Report 2020."

[21] Ivan Mehta, "[Best of 2019] WhatsApp Finally Lets You Prevent People from Adding You to Their Shitty Groups," *The Next Web* (blog), April 3, 2019, https://thenextweb.com/apps/2019/04/03/whatsapp-finally-lets-you-prevent-people-from-adding-you-to-their-shitty-groups/; Chinmayi Arun, "Opinion | India May Be Witnessing the next 'WhatsApp Election' — and the Stakes Couldn't Be Higher," *Washington Post*, April 25, 2019, https://www.washingtonpost.com/opinions/2019/04/25/india-could-see-next-whatsapp-election-stakes-couldnt-be-higher/; Priyanjana Bengani, "India Had Its First 'WhatsApp Election.' We Have a Million Messages from It," *Columbia Journalism Review* (blog), October 16, 2019, https://www.cjr.org/tow_center/india-whatsapp-analysis-election-security.php.

[22] Vindu Goel, "In India, Facebook's WhatsApp Plays Central Role in Elections," *The New York Times*, May 16, 2018, sec. Technology, https://www.nytimes.com/2018/05/14/technology/whatsapp-india-elections.html.

[23] Goel.

[24] Camila Mont'Alverne and Isabele Mitozo, "Muito Além da Mamadeira Erótica: As notícias compartilhadas nas redes de apoio a presidenciáveis em grupos de WhatsApp, nas eleições brasileiras de 2018 (Beyond the Erotic Baby Bottle: News shared through WhatsApp by presidential candidates' support networks during the Brazilian elections

outward bound YouTube links were for "The Antagonist," a political news media group highly favorable to current President Jair Bolsonaro.[25] In comparison, President Bolsonaro's opponents such as Fernando Haddad had visibly fewer mentions in these randomly sampled WhatsApp groups.

Closed messaging communications might in fact also help to carry an election. Beyond the anecdotal expressions by WhatsApp participants, the impact of the Messenger App influence appears to have worked in Brazil. Jair Bolsonaro — a candidate who had only 8 seconds in traditional political television campaign time — won over 55% of the electoral vote in October 2018.[26] How did this come to pass? Ten days before the 2018 election, Brazil's news publication Folha de São Paulo seemed to hint at a possible answer. Folha de São Paulo uncovered a scheme involving millions of campaign messages sent via WhatsApp in favor of Bolsonaro, all of which had been financed by various corporations affiliated to him.[27]

Finally, the misinformation flowing in these chats can foment life threatening behaviors. Already by 2017, WhatsApp officials in India were becoming concerned about Terms of Service violations including hate speech, threats of violence, and false statements.[28] False rumors within these chat spaces continued to incite extreme civil unrest, including cases of mob lynching, and contributed to anti-vaccination sentiment.[29] Due to concerns of rampant misinformation, WhatsApp began in 2018 to lower the impact of message reach by introducing limitations on forwards, focusing first on India "where people forward more messages, photos, and videos than any other country in the world."[30]

The conundrum of closed message space research and current practice

Closed messaging applications like WhatsApp have been used as a political mechanism to influence elections and to undermine the process through violence and civil unrest. However, these technologies are technically, or technologically, private. Whatever their public impact, these private spaces have been intentionally created to thwart hackers and company engineers, in addition to the general public.

This has created a certain awkwardness in the technology. For example, this year, amid the COVID-19 misinformation concerns, WhatsApp has continued to emphasize their efforts to

---

in 2018)," in *Compolítica8* (Compolítica8, Brasília, DF, 2019), 25, http://ctpol.unb.br/compolitica2019/GT4/gt4_Montalverne_Mitozo.pdf.

[25] Mont'Alverne and Mitozo.

[26] Esther Solano Gallego, "How Bolsonaro Came to Pass," *Rosa Luxemburg Stiftung NYC* (blog), November 27, 2018, http://www.rosalux-nyc.org/how-bolsonaro-came-to-pass/.

[27] Patrícia Campos Mello, "Empresários bancam campanha contra o PT pelo WhatsApp," *Folha de S.Paulo*, October 18, 2018, sec. Poder, https://web.archive.org/save/https://www1.folha.uol.com.br/poder/2018/10/empresarios-bancam-campanha-contra-o-pt-pelo-whatsapp.shtml, https://www1.folha.uol.com.br/poder/2018/10/empresarios-bancam-campanha-contra-o-pt-pelo-whatsapp.shtml.

[28] Goel, "In India, Facebook's WhatsApp Plays Central Role in Elections."

[29] Soutik Biswas, "Fighting India's WhatsApp Fake News War," *BBC News*, August 20, 2018, sec. India, https://www.bbc.com/news/world-asia-india-45140158. Referenced in Bengani, "India Had Its First 'WhatsApp Election.' We Have a Million Messages from It."

[30] WhatsApp Blog, "More Changes to Forwarding," *WhatsApp Blog* (blog), January 21, 2019, https://blog.whatsapp.com/more-changes-to-forwarding.

"constrain virality" and to signal "less personal" forwarded messages through the double arrow icon. With measures like these, the threat, at least according to WhatsApp, appears to be an external threat, not an internal one. Thus, the goal from the application's perspective is to keep the space "personal," "private," and "secure" – both in terms of encryption as well as misinformation.[31]

As a case in point, ownership of these messages appears to belong to end users; users are not contributing them to the public domain:

- According to WhatsApp Terms of Service: "Your messages are yours, and we can't read them. We've built privacy, end-to-end encryption, and other security features into WhatsApp. We don't store your messages once they've been delivered. When they are end-to-end encrypted, we and third parties can't read them."[32]
- Or, as the Telegram Privacy Policy emphasizes, "we don't use your data to show you ads," since data belongs to the end user; in addition, private chats are differentiated from public ones.[33]

Even application owners and developers cannot read the messages – although they are privy to quite a lot of metadata – because of encryption mechanisms. Therefore, in order to study what messages are flowing through these channels, analysts and researchers must join these private chat groups in ways that raise questions regarding the boundaries of ethical research in these contexts.

Regardless of the technological design for privacy within the chat space, these technologies are also at the same time designed to be *social* — they allow their users to open these spaces up to more public visibility and impact. We should acknowledge that the ways that the line between public and private conversation within these applications is explicitly blurred. For one, the scale of app reach – including a default technological capacity of well-over 200 members in the popular instance of WhatsApp or even 100,000 for Telegram – can at in groups with large membership defy any reasonable expectation of intimate conversation.

Secondly, group administrators and users sometimes publish invitations to chat groups through publicizing mechanisms that permit sharing: invitations to groups can be posted as a URL on websites, or in Instagram and Pinterest postings, for example. As another example of intersecting the private and public notions of closed chat spaces, Viber paired with an Israeli news station in 2015 to host a public moderated discussion on running for the Knesset.[34] Should a group decide to post an invitation to a group publicly, it seems logical to assume that the conversation may no longer be considered private.

Noting the sharing mechanism or the public availability of private chat links has not however completely resolved the question of user expectations. In February 2020, journalists reported that Google Search had been indexing WhatsApp Group invite links and rendering them publicly discoverable. Initially, responses from both platforms revealed an unsurprised posture; according

---

[31] WhatsApp Blog, "Keeping WhatsApp Personal and Private," *WhatsApp Blog* (blog), April 7, 2020, https://blog.whatsapp.com/Keeping-WhatsApp-Personal-and-Private.

[32] "WhatsApp Legal Info," WhatsApp.com, accessed July 8, 2020, https://www.whatsapp.com/legal/#key-updates.

[33] "Terms of Service," Telegram, accessed July 8, 2020, https://telegram.org/tos.

[34] Haleva-Amir, "Not All about That Facebook."

to Facebook, "like all content that is shared in searchable public channels, invite links that are posted publicly on the internet can be found by other WhatsApp users" and according to a Google representative, "Search engines like Google & others list pages from the open web. That's what's happening here. It's no different than any case where a site allows URLs to be publicly listed."[35] Despite claiming a no-fault position, Google took steps within a few days to remove the indexing and Facebook/WhatsApp supplied a no-follow rule to discourage automatic indexing as well.[36] As reported by articles on this topic showed, the sensibility of private and public discussion was central; "WhatsApp groups may not be as secure as you think they are," warned the Deutsche Welle reporter who initially broke the story.[37]

The technology therefore presents a conundrum for researchers, whether on the public-interest side or the academic side: these are technologies with large public effects but they take place within applications that are designed to be mostly private, and this complication poses difficulties for professional ethics.

Given the tension between public and private expectations in chat spaces, researchers have tried to mitigate the complication when researching closed messages in different ways. First, many of the studies reviewed have followed a growing practice, that they either do not collect or strive to de-identify phone number accounts from their data collections. For example, one study revealed phone numbers only at the country code level.[38] Studies can also opt not to analyze any user information, but focus instead just on the "public" topical content shared such as URLs or visual media considered to be part of the public domain.[39] Finally, some research efforts reported special data storage handling, such as the encryption of text messages or only using local hard drives.[40]

In addition, some studies sidestep the issue of messaging privacy altogether through design, by focusing only on what can be shared through an umbrella of informed consent. For example, studies can analyze app metadata in combination with participant questionnaires. Traditional methods of field study, specifically surveys and interviewing, continue to offer an avenue that

---

[35] Kim Lyons, "Google Is Indexing WhatsApp Group Chat Links, Making Even Private Groups Discoverable," *The Verge*, February 21, 2020, https://www.theverge.com/2020/2/21/21147073/whatsapp-google-group-chat-join-indexing-links-search-privacy-facebook.

[36] John Says, "Google Stops Indexing WhatsApp Chats; Other Search Engines Still at It," *Naked Security* (blog), February 25, 2020, https://nakedsecurity.sophos.com/2020/02/25/google-stops-indexing-whatsapp-chats-other-search-engines-still-at-it/.

[37] Joseph Cox, "VICE - Google Is Letting People Find Invites to Some Private WhatsApp Groups," *Motherboard (Vice)*, February 21, 2020, https://www.vice.com/en_us/article/k7enqn/google-is-letting-people-find-invites-to-some-private-whatsapp-groups.

[38] See for example Kiran Garimella and Gareth Tyson, "WhatsApp, Doc? A First Look at WhatsApp Public Group Data," in *Proceedings of the Twelfth International AAAI Conference on Web and Social Media* (International AAAI Conference on Web and Social Media, Stanford, CA: AAAI Press, 2018), https://aaai.org/Library/ICWSM/icwsm18contents.php; Josemar Alves Caetano et al., "Analyzing and Characterizing Political Discussions in WhatsApp Public Groups," *ArXiv:1804.00397 [Cs]*, April 2, 2018, http://arxiv.org/abs/1804.00397.

[39] Exemplified in Vidya Narayanan et al., "News and Information over Facebook and WhatsApp during the Indian Election Campaign," Computational Propaganda Project Data Memo (Oxford, UK: Oxford Internet Institute, May 13, 2019), https://comprop.oii.ox.ac.uk/research/india-election-memo/.

[40] See Garimella and Tyson, "WhatsApp, Doc?"; Bengani, "India Had Its First 'WhatsApp Election.' We Have a Million Messages from It."

can help characterize political behaviors around an election – such as generational trends or other aspects of community —without divulging specific examples of messages.[41] To be sure, these methods rely on the ability of participants to answer questions honestly and accurately as well as the capacity of investigators to appropriately contextualize the results.

However, in terms of understanding the actual discussions taking place within these spaces, there are studies that now examine message texts themselves. At least four models of research appear to be in operation — voluntary contribution, collection through focused partnerships, entrance with announcement or identification, and entrance without identification — which are explained further below.[42] They all define the public nature of their research, though in different ways.

Model 1: Do not enter private chat groups and rely on voluntary contributions (eg "tip lines" and broadcast messages).

A first model does not enter into chats and instead receives message texts from the contributions of application users, through an exchange that indicates consent.

Two cases demonstrate different ways that the "tip line" approach has been implemented:
- The Comprova project during the 2018 Elections in Brazil consisted of a coalition of twenty-four Brazilian newsrooms, backed by the First Draft Network. The coalition received tips of potential misinformation on WhatsApp including the sharing of message texts and images that an assigned journalist would then research and report back to the submitter and to the public.[43] In this manner, a member of the WhatsApp group could voluntarily submit a claim taking place within a conversation without revealing the identities of other group members or other aspects of the discussion.
- WhatsApp directly commissioned a similar model of tip line investigation during the 2019 Indian elections through the Checkpoint project, led by the Proto enterprise in partnership with Dig Deeper Media and Meedan companies. An important difference in this project model was that it was oriented more towards research for analysts than

---

[41] Maria Frahm-Arp, "The Political Rhetoric in Sermons and Select Social Media in Three Pentecostal Charismatic Evangelical Churches Leading up to the 2014 South African Election," *Journal for the Study of Religion* 28, no. 1 (2015): 115–41; Doris Ngozi Morah and Chinwe Elizabeth Uzochukwu, "Nigeria's Social Media Culture: Exploring Civic Participation of Youths in the 2015 Presidential Election," January 9, 2019, https://doi.org/10.5281/ZENODO.2536136; Erond L. Damanik, "Middle Class, WhatsApp, and Political Orientation: The Election of North Sumatera Governor, 2018" (Atlantis Press, 2019), 49–57, https://doi.org/10.2991/icssis-18.2019.11; Homero Gil de Zúñiga, Alberto Ardèvol-Abreu, and Andreu Casero-Ripollés, "WhatsApp Political Discussion, Conventional Participation and Activism: Exploring Direct, Indirect and Generational Effects," *Information Communication and Society*, July 20, 2019, 1–18, https://doi.org/10.1080/1369118X.2019.1642933.

[42] Examples of studies can be found in Bruno Schaefer et al., "Qual o Impacto Do WhatsApp Em Eleições? Uma Revisão Sistemática (2010-2019)," *Revista Debates* 13 (December 16, 2019): 58–88, https://doi.org/10.22456/1982-5269.96255.

[43] Anna Jean Kaiser, "The Brazilian Group Scanning WhatsApp for Disinformation in Run-up to Elections," *The Guardian*, September 26, 2018, sec. World news, https://www.theguardian.com/world/2018/sep/26/brazil-elections-comprova-project-misiniformation-whatsapp; Claire Wardle et al., "Comprova: An Evaluation of the Impact of a Collaborative Journalism Project on Brazilian Journalists and Audiences" (First Draft, June 27, 2019), https://firstdraftnews.org:443/latest/comprova-an-evaluation-of-the-impact-of-a-collaborative-journalism-project-on-brazilian-journalists-and-audiences/.

intended as a helpline for the public, which meant that not every submission received a response.[44]

As opposed to analyzing messages collected through tip lines, or many voluntary submissions to one single number (M:1 ratio), a third case demonstrates an analysis of messages available from one-to-many broadcasts (1:M):

- Researchers during the 2015 Presidential Elections in Nigeria analyzed WhatsApp broadcast messages. WhatsApp Broadcasts are sent from a sender to recipients on a defined list but can only be received if the recipient has stored the sender's phone number as well, presumably to prevent spamming.[45] By analyzing the content sent by select accounts to many recipients —a mechanism of implicit consent through the mutual storage of phone numbers —the researchers were able to give a sense of the tone, accusations, and assertions around political candidates that were part of the election media environment without entering into private chat situations or seeing other's responses.

Model 2: Enter specific chat groups with invitation or consent for a publicly identified purpose.

Another model for evaluating messages, but this time with members of the analysis team entering directly into chat spaces themselves, was established through a collaborative election tracking project. Unlike Model 1, this project collected all messages from the closed group for a certain period, which meant that analysts could examine both the text of a message and the account details of the message sender, in addition to the conversational context:

- During the Ghana 2016 general election, a collaboration of organizations established a "Social Media Tracking Center" (SMTC) in order to prevent violence and other election threats by monitoring messages on social media.[46] The Tracking Center model and software platform had previously been established in other country election contexts to examine Twitter and Facebook information but also, for the Ghana election, included the collection of WhatsApp messages. In advance of the election, members of Pen Plus Bytes (a local civil society organization), United Nations University, Ghana's National Election Security Task Force, the Electoral Commission among others established protocols of communication for verified incidents.

  A single SMTC WhatsApp account, which was publicly shared, meant that citizens could submit their own questions and reports. In addition, this account was permitted access by

---

[44] Abhimanyu Ghoshal, "WhatsApp Launches a Tip Line in India to Battle Fake News," *The Next Web*, April 2, 2019, https://thenextweb.com/apps/2019/04/02/whatsapp-launches-a-tip-line-in-india-to-battle-fake-news-ahead-of-national-elections/; Ryan Mac and Pranav Dixit, "The Tip Line WhatsApp Launched to Combat Fake News Isn't Actually Going to Combat Fake News," *BuzzFeed News*, April 3, 2019, https://www.buzzfeednews.com/article/ryanmac/whatsapp-fake-news-tip-line-indian-election-not-helpline; "PROTO | Checkpoint 'FAQs About Project Checkpoint,'" Proto, accessed July 8, 2020, https://www.checkpoint.pro.to/.

[45] Oluwabunmi O. Oyebode and Adeyemi Adegoju, "Appraisal Resources in Select WhatsApp Political Broadcast Messages in the 2015 Presidential Election Campaign in Nigeria," *Journal of Pan African Studies* 10, no. 10 (November 1, 2017): 29; "WhatsApp FAQ - How to Use Broadcast Lists," WhatsApp.com, accessed July 13, 2020, https://faq.whatsapp.com/iphone/chats/how-to-use-broadcast-lists.

[46] Andrés Moreno, Philip Garrison, and Karthik Bhat, "WhatsApp for Monitoring and Response during Critical Events: Aggie in the Ghana 2016 Election," in *Proceedings of the 14th ISCRAM Conference* (ISCRAM 2017, Albi, France, 2017), 11.

group administrators to a set of WhatsApp groups for incident tracking. The texts of all messages were collected — something explained to the members of the WhatsApp group by Pen Plus Bytes and the research team. This fuller collection allowed the SMTC software to automatically flag potential issues for further examination by a team of human reviewers who could escalate issues when necessary.[47]

Model 3: Enter "public" chat groups with research identification, allowing for removal or withdrawal when requested.

The third model leverages the ambiguity of private chat invitations, or links available on the open web, in order to enter closed groups and study them. Several examples from recent elections demonstrate this practice in slightly different ways, all of which included some form of identification as a researcher within the chat space.

- In one approach, researchers use the "name" field of the account to identify themselves, but do not make declarations beyond this. Kiran Garimella and Gareth Tyson designed a methodology and software process to scrape what they defined as public WhatsApp data at a larger scale than typical studies. The authors took chat messages from a sample of 200 groups from 2,500 WhatsApp groups in 2018 that were publicly indexed.[48] When joining such groups, the researchers identified the accounts as being "researchers from" a research organization, but placed no announcement within the group chat about their joining or presence as researchers.[49] This methodology and process has now been systematized further by the WhatsApp Monitor system project.[50]
- Another example of this approach —identification but no announcement – took place during the 2019 India elections.[51] The collection of over one million messages by the Tow Center for Journalism at Columbia University for the period between April 11 and May 19, 2019 leveraged both "groups of which we were already a part" as well as joined other groups with publicized invite links of at least 60 members. The study noted that they were at times banned from WhatsApp or dropped from groups but that the rationale for removal was not clear (e.g., automated group joins from WhatsApp Desktop potentially flagged as bots, a foreign US-based account number flagged to administrators).[52]
- In a different case from the Indian 2019 elections, researchers from Oxford University identified themselves in two ways. After finding links to chat groups index on the web that fit their parameters, they announced their identity upon gaining entry to the chat

---

[47] According to the project, there was an "introductory text message that was sent [by members of Pen Plus Bytes and the research team] from the SMTC phone to the WhatsApp groups being monitored. It briefly explained what the SMTC was and that all the messages in the group were being scooped up," Philip Garrison to Connie Moon Sehat, "[Request for Clarification] Research Practices for Closed Messaging Groups," July 22, 2020.

[48] Garimella and Tyson, "WhatsApp, Doc?"

[49] Kiran Garimella to Connie Moon Sehat, "Re: [Air-L] [Request for Input] Research Practices for Closed Messaging Groups," March 25, 2020.

[50] Philipe Melo et al., "WhatsApp Monitor: A Fact-Checking System for WhatsApp," *Proceedings of the International AAAI Conference on Web and Social Media* 13 (July 6, 2019): 676–77.

[51] "to be fully transparent, we identified ourselves as 'Tow Center' with the description 'a research group based out of Columbia University in New York,'" from Priyanjana Bengani to Aleksei Kaminski, "Email Reply: Introducing Myself and Your Research," March 18, 2020.

[52] Bengani, "India Had Its First 'WhatsApp Election.' We Have a Million Messages from It."

space. The purpose was to offer "group members the option to withdraw consent from participating in this study. This resulted in some administrators removing us."[53]

In contrast to most of the cases from the first models, research results from Model 3 were not directly communicated back to chat group members or to civil society and government officials; rather, their output was for academic purposes of understanding. Their research designs, sometimes approved by university review boards, may not have required the researchers to report their findings back to the members of groups they joined.

Model 4: Enter more "public" private groups but without any identification.

Finally, additional cases demonstrate how researchers have sometimes leveraged the ambiguity of the publicly indexed chat groups without disclosing their researcher identities to group participants.

This model raises the question of whether the notion of a "public WhatsApp group" – when publicized and made available for general entry – does exist generally among the public or perhaps for a specific context (eg. topic, cultural) that lowers the barrier of entry for research purposes.

- In the Mont'Alverne and Mitozo example from Brazil 2018 cited earlier, the researchers did not share their purposes.[54] According to one of the researchers, they felt that divulging their presence would change the dynamics of the conversation and would likely result in further exclusions from groups; however, they stressed a protocol that strove to preserve identities in the research. Having collected their WhatsApp invitation links from open social networking locations on Twitter and Facebook in addition to Google, the researchers also stressed that they considered these to be public arenas and open for entry.[55]

Other cases from Brazil in 2018 and 2019 from researchers at the Pontifícia Universidade Católica de Minas Gerais (PUC Minas) and Universidade Federal de Minas Gerais (UFMG) – in addition to IBM Research and the Max Planck Institute for Software Systems (MPI-SWS) in Germany —also appear in part to have followed the same protocol of not indicating researcher presence.[56]

---

[53] Narayanan et al., "News and Information over Facebook and WhatsApp during the Indian Election Campaign"; Vidya Narayanan to Connie Moon Sehat, "Re: [Request for Input] Research Practices for Closed Messaging Groups," July 28, 2020.

[54] Mont'Alverne and Mitozo, "Muito Além da Mamadeira Erótica: As notícias compartilhadas nas redes de apoio a presidenciáveis em grupos de WhatsApp, nas eleições brasileiras de 2018 (Beyond the Erotic Baby Bottle: News shared through WhatsApp by presidential candidates' support networks during the Brazilian elections in 2018)."

[55] « Nós fizemos a opção de não nos identificar ao entrar nos grupos. Isso aconteceu porque não haveria outra forma de realizar a pesquisa. Os grupos tendem a ser muito fechados e nos identificar como pesquisadores alteraria completamente a dinâmica e, em outros casos, seríamos excluídos (já fomos excluídos de alguns sem mesmo nos identificar), » Camila Mont'Alverne to Aleksei Kaminski, "E-Mail Reply: Introdução e Sua Pesquisa-Aleksei Kaminski," March 18, 2020; Camila Mont'Alverne to Aleksei Kaminski and Connie Moon Sehat, "Re: Oi e Conselho! - Aleksei," July 23, 2020.

[56] Other examples include Caetano et al., "Analyzing and Characterizing Political Discussions in WhatsApp Public Groups"; Gustavo Resende et al., "Analyzing Textual (Mis)Information Shared in WhatsApp Groups," in

Model 3 and Model 4 can be blended; a joint project among researchers from institutions with different requirements can result in this mix. For example, for a project across MIT, MPI-SWS, and other institutions, the part of the data collection that took place under MPI-SWS did not have to disclose researcher status.[57]

Questions for Researchers, Questions for the Public

Through practices that seek to protect the identities of chat participants, we find that many investigators are clearly trying to follow ethics established by their various professional communities. They do this by designing approaches that consider possible, future harm upon those observed. In recent years, expectations and requirements related to personal data collection and storage online have risen overall, making this obligation clearer for researchers across the board as well.[58]

The challenge however for closed messaging is that both the technology and its ethics are still emerging.[59] At bottom, we lack a scaled understanding or approach to the issues posed by closed messaging: due to the barriers for access or entry, current studies of messages flowing through these spaces tend to remain very small and are necessarily limited in terms of their ability to speak to larger trends. How severe is the problem of misinformation or disinformation within closed message chats, for example? Depending on how the application is built, even the application owners do not know the answer to this question, since the encryption layer creates a hurdle that even internal engineers are unable to cross without trying to break the system.

Still, the technology allows for scaling, far beyond the intimacy that encrypted chat suggests. In order to begin to answer the academic question of larger impact, approaches offered in Model 3 and 4 intentionally try to move towards conclusions that might be more generalizable about the "public."[60] A yet further complication to this kind of public research is the potential variance in country or cultural expectations across *multiple publics,* let alone from different group participants, in terms of research taking place within their chats. Perhaps, in other words, the expectations for privacy or the public-private divide is different according to topic, culture, and country.

Proceedings of the 10th ACM Conference on Web Science, WebSci '19 (Boston, Massachusetts, USA: Association for Computing Machinery, 2019), 225–234, https://doi.org/10.1145/3292522.3326029; Gustavo Resende et al., "(Mis)Information Dissemination in WhatsApp: Gathering, Analyzing and Countermeasures," in *The World Wide Web Conference on - WWW '19* (The World Wide Web Conference, San Francisco, CA, USA: ACM Press, 2019), 818–28, https://doi.org/10.1145/3308558.3313688.

[57] Fabricio Benevenuto to Aleksei Kaminski and Connie Moon Sehat, "Introducing Myself and Your Research-Aleksei," September 3, 2020.

[58] Such as the Common Rule in the United States (add citation) or European Commission and Directorate General for Research, *Ethics for Researchers: Facilitating Research Excellence in FP7.* (Luxembourg: Publications Office, 2013), http://bookshop.europa.eu/uri?target=EUB:NOTICE:KI3213114:EN:HTML.

[59] See for example a discussion on citizen journalism understanding of ethics in Shepherd Mpofu and Shanade Bianca Barnabas, "Citizen Journalism and Moral Panics: A Consideration of Ethics in the 2015 South African Xenophobic Attacks," *African Journalism Studies* 37, no. 4 (November 2016): 115–36, https://doi.org/10.1080/23743670.2016.1256053. n

[60] Halford et al., "Understanding the Production and Circulation of Social Media Data."

Despite the potential for different expectations among differing publics, there do appear to be a number of questions that researchers are either implicitly or explicitly deciding before embarking upon the collection and analysis of closed message texts in democratic contexts:

- **Exactly when is a closed message chat "public"?** [What are the conditions in terms of indexed invites, group size, discussion topics, or other aspects of closed message discussions that make them arguably public versus private?]

- **Who does the data belong to?** And how should it be stored or shared?

- **What are the obligations for a) researcher disclosure and or b) informed consent?** (especially given the ongoing nature of chat group membership)

- **When should researchers to inform or report back to the groups involved the findings of their studies**? And how?

In addition, two questions are suggested by our landscape review:
- **Are these questions researchers can ask the public (or publics)?**
- **Are these questions that researchers should discuss with companies**?

All these questions are defined in further detail below.

1. *Exactly when is a closed message chat "public"? [What are the conditions in terms of indexed invites, group size, discussion topics, or other aspects of closed message discussions that make them arguably public versus private?]*

Publicly indexed/available invite? The answer to this question may become less (or more?) pressing given the February removal of WhatsApp URL indexing from Google, mentioned above, but the question remains: does a publicized invitation render a closed message space conversation as "public"? In the European context, for example, the notion of the "right to respect for his or her family life, home and communications" is strong, so understanding this boundary is important.[61]

First Draft's approach in 2018 explained that their tip line approach seemed to be the only way to honor the Terms of Service: "Due to WhatsApp's end-to-end encryption, this method of soliciting tips is the only possible way to collect misinformation data without violating the app's terms of service."[62]

Topic and scope? At the same time, the need to understand the broad impact of social media technologies is contributing to new interpretations of "fair use" in service of the public interest. In the United States, recent rulings have sanctioned Terms of Service violations when researching bias and racial discrimination in online hiring websites. The creation of false logins or identities, a breach of website terms of service, has recently been judged *not* to be a violation

---

[61] Article 7, "Charter of Fundamental Rights of the European Union," *Official Journal of the European Union*, no. C 83 (March 30, 2010): 389–403.

[62] Pedro Burgos, "What 100,000 WhatsApp Messages Reveal about Misinformation in Brazil," *First Draft* (blog), June 27, 2019, https://firstdraftnews.org:443/latest/what-100000-whatsapp-messages-reveal-about-misinformation-in-brazil/.

of the Computer Fraud and Abuse Act for certain research purposes.[63] Upcoming cases will also continue to wrestle with the legality of web scraping on arguably public, or at least, publicly available LinkedIn profile pages.[64]

Group size and expectation? Finally, size also factors in a definition of "public" conversation. The expectation of privacy, or at least a private, intimate conversation, among a 100,000- or even 200-member chat may be low, no matter the topic. It may be that a group conversation of 5 or 10, even if with a publicly indexed link on a topic of public interest, still has an appropriate expectation of privacy. The Columbia study used a threshold of 60 members as crossing over into the realm of public analysis.

2. **Who does the data belong to?** *And how should it be stored or shared?*

Defining who owns the data has implications for how research is conducted and how the data is stored or shared. As discussed earlier, message texts belong to the message writer, at least according to two examples of Terms of Service. However, the recent *Sandvig v. Barr* decision indicates the possible allowance of researching such information under certain conditions and in the public interest.

Furthermore, from the Oxford Internet Institute example of message analysis (under Model 3), we see a practice in which the researchers employed a content-oriented strategy. They focused not on personal texts but generally accessible media such as URLs and images shared through in the information network. Perhaps in this case, if focusing on a particular element of data, the considerations are different.

Data related to the message writers and receivers, such as phone numbers, belong also to the account holders. In aggregate, however, to whom does the data belong, and in what cases may it be shared? The European Union's General Data Protection Regulation (GDPR) has the strongest considerations of what is considered personal data, which at times includes the IP addresses through which citizens are active.[65] Yet when a citizen voluntarily enters a public arena, some identification is necessarily divulged; defining the context for communication is important.

3. **What are the obligations for a) researcher disclosure and or b) informed consent?**
   (especially given the ongoing nature of chat group membership)

The chat space does not lend itself well for traditional practices of researcher disclosure and informed consent not just because of its encrypted nature. In addition, the ongoing inclusions of new members and departures of older ones creates a challenge. Scholars Barbosa and Milan have argued, "even when one discloses their researcher identity, the fluidity of belonging and participation in a chat app creates a situation that it is not really possible to inform every group

---

[63] Sandvig v. Barr, No. 16–1368 (D.D.C. March 27, 2020).

[64] hiQ Labs, Inc. v. LinkedIn Corporation, No. 17–16783 (9th Cir. September 9, 2019).

[65] Eg, European Union General Data Protection Regulation, Article 3, European Union, "General Data Protection Regulation (GDPR)," REGULATION (EU) 2016/ 679 § Art. 4 GDPR - Definitions (2018), https://gdpr.eu/article-4-definitions/.

participant about the ongoing observation, let alone remind each of it; ensuring 'informed' consent is probably just wishful thinking in very large groups."[66]

However, researcher disclosure and informed consent may not be required by different professional practices for all types of investigation. In the United States for example, oral history, journalism, literary criticism, legal research, and historical scholarship, are not required to practice informed consent because they do not contribute to generalizable knowledge but instead specific knowledge of certain individuals.[67] The observation of public behavior, such as through election observation, can also fall under this exemption.[68] Practitioners of these fields may still self-regulate, and define best practices and ethical codes that include informed consent.[69] In the European Union, in the wake of Nazi-era experimentation, free and informed consent has been emphasized in the context of medicine and biological research within the Nuremberg Codes as well as the European Charter of Fundamental Rights.[70]

As final note, answering the question of who owns the data (Question 2) may have implications for the answer to this question.

4. *When should researchers to inform or report back to the groups involved the findings of their studies*? *And how?*

This question is more relevant for Models 3 and 4, or projects that focus on academic outcomes, rather than those that report directly back to citizens as part of their design, such as with election incident reporting.

Question 4 considers the traditions from which informed consent emerged. For example, take the case of the Tuskegee Study in the United States, which demonstrated necessary limits within democracies engaging in human subject research upon their own citizens.

Beginning in 1932, the US Public Health Service conducted the "Tuskegee Study of Untreated Syphilis in the Negro Male" for 40 years, which amounted to a passive observation of the effects of syphilis in hundreds of African-American men.[71] In doing so, the US Public Health Service misled hundreds of participants, and neglected to inform them of proper treatment options for the disease once their effectiveness became established.

---

[66] Barbosa and Milan, "Do Not Harm in Private Chat Apps."

[67] Indeed, in this definition, the activities are not considered "research," which the US Federal Register only limits to activities of a systematic, generalizable nature. "Federal Policy for the Protection of Human Subjects," *Federal Register (United States)*, Rules and Regulations, 82, no. 12 (January 19, 2017): 7149-.

[68] For one example, as defined in Emory University, "Emory University Institutional Review Board Policies and Procedures," February 11, 2020, http://www.irb.emory.edu/documents/PoliciesandProcedures.pdf.

[69] As for example, the Oral History Association, as referenced by American Historical Association, "Public Statement on Oral History and Human Subjects Regulation (November 2015)" (American Historical Association, November 12, 2015), https://www.historians.org/news-and-advocacy/aha-advocacy/public-statement-on-oral-history-and-human-subjects-regulation.

[70] "Charter of Fundamental Rights of the European Union."

[71] First captured October 11, 2007 at https://web.archive.org/web/20071011014512/http://www.cdc.gov/tuskegee/timeline.htm, "The U.S. Public Health Service Syphilis Study at Tuskegee: The Tuskegee Timeline," Centers for Disease Control and Prevention, 2007 2020, http://www.cdc.gov/tuskegee/timeline.htm.

Especially when the topic of exchange is not merely misleading but potentially harmful – one of the motivations of this research outlined earlier – do researchers have an obligation to report their results back in certain cases? If so, what are the thresholds?

Again, the question of who owns the data (Question 2) may have implications for this area.

5. *Are these questions researchers can ask the public (or publics)?* *Are there ways to understand the expectations of the community, cultural, or country context?*
6. *Are these questions that researchers should discuss with companies?*

Given that closed message research deals with the public, it seems worth asking whether or not questions about private versus public conversations might be asked directly to representatives of the communities or countries involved. A different set of researchers might need to undertake this inquiry in order to mitigate any conflict of interest. As the effort is likely large in itself, in order to be representative, other methods might be more useful – surveys might be employed. The effort to gain public feedback may also require different relationships between researchers and the public, different obligations of reporting, and different research approaches. Yet obtaining answers to Question 5 may offer an opportunity for collaborating across different organizations or disciplines, in order to establish larger consensus related to studies taking place in closed messaging spaces.

A key issue when it comes to closed message research is that most approaches do not have direct access to app data (partnership with corporate research efforts may offer exceptions). In the case of closed messaging applications, the potential for an exchange among companies that own these applications, researchers, and the public is inherently limited; encryption makes this so. Furthermore, self-reported or company-provided data is not the same as externally audited information.

Given the ways that these spaces are by design both public and private – at times one more than the other – an improved exchange among companies, publics, and researchers may lead to creative approaches and answers to all the questions outlined above. Perhaps, for example, greater clarity about the boundaries between private and public space could be jointly discussed and communicated.

Next Steps

As mentioned at the outset, past workshops in 2018 and 2019 laid the foundation for this paper's inquiry. Calls for suggestions of related case studies, as well as for researchers who might be interested in collaborating, were issued to the same workshop participants in addition to the Association of Internet Researchers and Responsible Data Forum group email lists in the first half of 2020. We greatly appreciate the clarifications regarding research practices that we received from many of the investigators whose studies were cited under the four models.

A second draft of the paper will follow a period of invited consultation in collaboration with The Carter Center, with a finalized version intended for the beginning of 2021. We invite suggestions and feedback to http://bit.ly/closed-message-research-feedback.

References

American Historical Association. "Public Statement on Oral History and Human Subjects Regulation (November 2015)." American Historical Association, November 12, 2015. https://www.historians.org/news-and-advocacy/aha-advocacy/public-statement-on-oral-history-and-human-subjects-regulation.

Anderson, Meghan Keaney. "Why We're Thinking About Messaging Apps All Wrong." *HubSpot* (blog), February 9, 2016. https://blog.hubspot.com/marketing/messaging-apps.

Arun, Chinmayi. "Opinion | India May Be Witnessing the next 'WhatsApp Election' — and the Stakes Couldn't Be Higher." *Washington Post*, April 25, 2019. https://www.washingtonpost.com/opinions/2019/04/25/india-could-see-next-whatsapp-election-stakes-couldnt-be-higher/.

Barbosa, Sérgio, and Stefania Milan. "Do Not Harm in Private Chat Apps: Ethical Issues for Research on and with WhatsApp." *Westminster Papers in Communication and Culture* 14, no. 1 (August 14, 2019): 49–65. https://doi.org/10.16997/wpcc.313.

Bengani, Priyanjana. Letter to Aleksei Kaminski. "Email Reply: Introducing Myself and Your Research," March 18, 2020.

———. "India Had Its First 'WhatsApp Election.' We Have a Million Messages from It." *Columbia Journalism Review* (blog), October 16, 2019. https://www.cjr.org/tow_center/india-whatsapp-analysis-election-security.php.

Biswas, Soutik. "Fighting India's WhatsApp Fake News War." *BBC News*, August 20, 2018, sec. India. https://www.bbc.com/news/world-asia-india-45140158.

Caetano, Josemar Alves, Jaqueline Faria de Oliveira, Helder Seixas Lima, Humberto T. Marques-Neto, Gabriel Magno, Wagner Meira Jr, and Virgílio A. F. Almeida. "Analyzing and Characterizing Political Discussions in WhatsApp Public Groups." *ArXiv:1804.00397 [Cs]*, April 2, 2018. http://arxiv.org/abs/1804.00397.

"Charter of Fundamental Rights of the European Union." *Official Journal of the European Union*, no. C 83 (March 30, 2010): 389–403.

Cox, Joseph. "VICE - Google Is Letting People Find Invites to Some Private WhatsApp Groups." *Motherboard (Vice)*, February 21, 2020. https://www.vice.com/en_us/article/k7enqn/google-is-letting-people-find-invites-to-some-private-whatsapp-groups.

Damanik, Erond L. "Middle Class, WhatsApp, and Political Orientation: The Election of North Sumatera Governor, 2018," 49–57. Atlantis Press, 2019. https://doi.org/10.2991/icssis-18.2019.11.

Dotto, Carlotta, Rory Smith, and Claire Wardle. "Closed Groups, Messaging Apps, and Online Ads." First Draft's Essential Guide. First Draft, November 2019. https://firstdraftnews.org/wp-content/uploads/2019/11/Messaging_Apps_Digital_AW-1.pdf?x80491.

Emory University. "Emory University Institutional Review Board Policies and Procedures," February 11, 2020. http://www.irb.emory.edu/documents/PoliciesandProcedures.pdf.

European Commission, and Directorate General for Research. *Ethics for Researchers: Facilitating Research Excellence in FP7.* Luxembourg: Publications Office, 2013. http://bookshop.europa.eu/uri?target=EUB:NOTICE:KI3213114:EN:HTML.

European Union. General Data Protection Regulation (GDPR), REGULATION (EU) 2016/ 679 § Art. 4 GDPR - Definitions (2018). https://gdpr.eu/article-4-definitions/.

Fabricio Benevenuto. Letter to Aleksei Kaminski and Connie Moon Sehat. "Introducing Myself and Your Research- Aleksei," September 3, 2020.

"Federal Policy for the Protection of Human Subjects." *Federal Register (United States)*, Rules and Regulations, 82, no. 12 (January 19, 2017): 7149-.

Frahm-Arp, Maria. "The Political Rhetoric in Sermons and Select Social Media in Three Pentecostal Charismatic Evangelical Churches Leading up to the 2014 South African Election." *Journal for the Study of Religion* 28, no. 1 (2015): 115–41.

Garimella, Kiran. Letter to Connie Moon Sehat. "Re: [Air-L] [Request for Input] Research Practices for Closed Messaging Groups," March 25, 2020.

Garimella, Kiran, and Gareth Tyson. "WhatsApp, Doc? A First Look at WhatsApp Public Group Data." In *Proceedings of the Twelfth International AAAI Conference on Web and Social Media*. Stanford, CA: AAAI Press, 2018. https://aaai.org/Library/ICWSM/icwsm18contents.php.

Garrison, Philip. Letter to Connie Moon Sehat. "[Request for Clarification] Research Practices for Closed Messaging Groups," July 22, 2020.

Ghoshal, Abhimanyu. "WhatsApp Launches a Tip Line in India to Battle Fake News." *The Next Web*, April 2, 2019. https://thenextweb.com/apps/2019/04/02/whatsapp-launches-a-tip-line-in-india-to-battle-fake-news-ahead-of-national-elections/.

Gil de Zúñiga, Homero, Alberto Ardèvol-Abreu, and Andreu Casero-Ripollés. "WhatsApp Political Discussion, Conventional Participation and Activism: Exploring Direct, Indirect and Generational Effects." *Information Communication and Society*, July 20, 2019, 1–18. https://doi.org/10.1080/1369118X.2019.1642933.

Goel, Vindu. "In India, Facebook's WhatsApp Plays Central Role in Elections." *The New York Times*, May 16, 2018, sec. Technology. https://www.nytimes.com/2018/05/14/technology/whatsapp-india-elections.html.

Haleva-Amir, Sharon. "Not All about That Facebook: Political Campaigns and Civic Engagement in the 2015 Elections." *Israel Affairs* 22, no. 3/4 (July 2016): 711–26. https://doi.org/10.1080/13537121.2016.1174375.

Halford, Susan, Mark Weal, Ramine Tinati, Les Carr, and Catherine Pope. "Understanding the Production and Circulation of Social Media Data: Towards Methodological Principles and Praxis." *New Media & Society* 20, no. 9 (September 2018): 3341–58. https://doi.org/10.1177/1461444817748953.

Hegde, Pranav. "Coronavirus Impact | WhatsApp Global Usage Spikes by 51% during Pandemic." *MSN News: Money Control* (blog), March 26, 2020. https://www.msn.com/en-in/news/other/coronavirus-impact-whatsapp-global-usage-spikes-by-51-25-during-pandemic/ar-BB11JFMt.

hiQ Labs, Inc. v. LinkedIn Corporation, No. 17–16783 (9th Cir. September 9, 2019).

International Telecommunication Union. "Statistics." *International Telecommunication Union* (blog). Accessed June 30, 2020. https://www.itu.int/en/ITU-D/Statistics/Pages/stat/default.aspx.

Kaiser, Anna Jean. "The Brazilian Group Scanning WhatsApp for Disinformation in Run-up to Elections." *The Guardian*, September 26, 2018, sec. World news. https://www.theguardian.com/world/2018/sep/26/brazil-elections-comprova-project-misiniformation-whatsapp.

Kemp, Simon. "Digital 2020: Global Digital Overview." DataReportal in collaboration with HootSuite and WeAreSocial, January 30, 2020. https://datareportal.com/reports/digital-2020-global-digital-overview.

Klosowski, Thorin. "What's the Deal with All These Messaging Apps?" *Lifehacker* (blog), April 10, 2014. https://lifehacker.com/whats-the-deal-with-all-these-messaging-apps-1561543034.

Korean Ministry of Science and ICT. "2018년 인터넷이용실태조사 결과 발표 (Announcement of the 2018 Internet Usage Survey)." *대한민국 정책포털KOREA.KR 정책브리핑 (Korea Policy Portal KOREA.KR Policy Briefing)* (blog), February 25, 2019. http://www.korea.kr/news/pressReleaseView.do?newsId=156318833.

Lyons, Kim. "Google Is Indexing WhatsApp Group Chat Links, Making Even Private Groups Discoverable." *The Verge*, February 21, 2020. https://www.theverge.com/2020/2/21/21147073/whatsapp-google-group-chat-join-indexing-links-search-privacy-facebook.

Mac, Ryan, and Pranav Dixit. "The Tip Line WhatsApp Launched to Combat Fake News Isn't Actually Going to Combat Fake News." *BuzzFeed News*, April 3, 2019. https://www.buzzfeednews.com/article/ryanmac/whatsapp-fake-news-tip-line-indian-election-not-helpline.

McMahon, Jordan. "Why We Should All Ditch Other Messaging Apps for Signal." *Wired*, November 5, 2017. https://www.wired.com/story/ditch-all-those-other-messaging-apps-heres-why-you-should-use-signal/.

Mehta, Ivan. "[Best of 2019] WhatsApp Finally Lets You Prevent People from Adding You to Their Shitty Groups." *The Next Web* (blog), April 3, 2019. https://thenextweb.com/apps/2019/04/03/whatsapp-finally-lets-you-prevent-people-from-adding-you-to-their-shitty-groups/.

Mello, Patrícia Campos. "Empresários bancam campanha contra o PT pelo WhatsApp." *Folha de S.Paulo*, October 18, 2018, sec. Poder. https://web.archive.org/save/https://www1.folha.uol.com.br/poder/2018/10/empresarios-bancam-campanha-contra-o-pt-pelo-whatsapp.shtml. https://www1.folha.uol.com.br/poder/2018/10/empresarios-bancam-campanha-contra-o-pt-pelo-whatsapp.shtml.

Melo, Philipe, Johnnatan Messias, Gustavo Resende, Kiran Garimella, Jussara Almeida, and Fabrício Benevenuto. "WhatsApp Monitor: A Fact-Checking System for WhatsApp." *Proceedings of the International AAAI Conference on Web and Social Media* 13 (July 6, 2019): 676–77.

"Mobile Services - Bots, the next Frontier." *The Economist*, April 9, 2016, sec. Business. https://www.economist.com/business/2016/04/09/bots-the-next-frontier.

Mont'Alverne, Camila. Letter to Aleksei Kaminski. "E-Mail Reply: Introdução e Sua Pesquisa-Aleksei Kaminski," March 18, 2020.

———. Letter to Aleksei Kaminski and Connie Moon Sehat. "Re: Oi e Conselho! - Aleksei," July 23, 2020.

Mont'Alverne, Camila, and Isabele Mitozo. "Muito Além da Mamadeira Erótica: As notícias compartilhadas nas redes de apoio a presidenciáveis em grupos de WhatsApp, nas eleições brasileiras de 2018 (Beyond the Erotic Baby Bottle: News shared through WhatsApp by presidential candidates' support networks during the Brazilian elections in 2018)." In *Compolítica8*, 25. Brasília, DF, 2019. http://ctpol.unb.br/compolitica2019/GT4/gt4_Montalverne_Mitozo.pdf.

Morah, Doris Ngozi, and Chinwe Elizabeth Uzochukwu. "Nigeria's Social Media Culture: Exploring Civic Participation of Youths in the 2015 Presidential Election," January 9, 2019. https://doi.org/10.5281/ZENODO.2536136.

Moreno, Andrés, Philip Garrison, and Karthik Bhat. "WhatsApp for Monitoring and Response during Critical Events: Aggie in the Ghana 2016 Election." In *Proceedings of the 14th ISCRAM Conference*, 11. Albi, France, 2017.

Mpofu, Shepherd, and Shanade Bianca Barnabas. "Citizen Journalism and Moral Panics: A Consideration of Ethics in the 2015 South African Xenophobic Attacks." *African Journalism Studies* 37, no. 4 (November 2016): 115–36. https://doi.org/10.1080/23743670.2016.1256053.

Narayanan, Vidya. Letter to Connie Moon Sehat. "Re: [Request for Input] Research Practices for Closed Messaging Groups," July 28, 2020.

Narayanan, Vidya, Bence Kollanyi, Ruchi Hajela, Ankita Barthwal, Nahema Marchal, and Philip N. Howard. "News and Information over Facebook and WhatsApp during the Indian Election Campaign." Computational Propaganda Project Data Memo. Oxford, UK: Oxford Internet Institute, May 13, 2019. https://comprop.oii.ox.ac.uk/research/india-election-memo/.

Newman, Nic. "Reuters Institute Digital News Report 2020." Reuters Institute for the Study of Journalism, 2020. https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2020-06/DNR_2020_FINAL.pdf.

Newman, Nic, Richard Fletcher, Antonis Kalogeropoulos, David A. L. Levy, and Rasmus Kleis Nielsen. "Reuters Institute Digital News Report 2018." Reuters Institute for the Study of Journalism, 2018.

Oyebode, Oluwabunmi O., and Adeyemi Adegoju. "Appraisal Resources in Select WhatsApp Political Broadcast Messages in the 2015 Presidential Election Campaign in Nigeria." *Journal of Pan African Studies* 10, no. 10 (November 1, 2017): 29.

Pedro Burgos. "What 100,000 WhatsApp Messages Reveal about Misinformation in Brazil." *First Draft* (blog), June 27, 2019. https://firstdraftnews.org:443/latest/what-100000-whatsapp-messages-reveal-about-misinformation-in-brazil/.

Proto. "PROTO | Checkpoint 'FAQs About Project Checkpoint.'" Accessed July 8, 2020. https://www.checkpoint.pro.to/.

Resende, Gustavo, Philipe Melo, Julio C. S. Reis, Marisa Vasconcelos, Jussara M. Almeida, and Fabrício Benevenuto. "Analyzing Textual (Mis)Information Shared in WhatsApp Groups." In *Proceedings of the 10th ACM Conference on Web Science*, 225–234. WebSci '19. Boston, Massachusetts, USA: Association for Computing Machinery, 2019. https://doi.org/10.1145/3292522.3326029.

Resende, Gustavo, Philipe Melo, Hugo Sousa, Johnnatan Messias, Marisa Vasconcelos, Jussara Almeida, and Fabrício Benevenuto. "(Mis)Information Dissemination in WhatsApp: Gathering, Analyzing and Countermeasures." In *The World Wide Web Conference on* -

WWW '19, 818–28. San Francisco, CA, USA: ACM Press, 2019. https://doi.org/10.1145/3308558.3313688.

Sandvig v. Barr, No. 16–1368 (D.D.C. March 27, 2020).

Says, John. "Google Stops Indexing WhatsApp Chats; Other Search Engines Still at It." *Naked Security* (blog), February 25, 2020. https://nakedsecurity.sophos.com/2020/02/25/google-stops-indexing-whatsapp-chats-other-search-engines-still-at-it/.

Schaefer, Bruno, Tiago Barbosa, Sara Epitácio, and Roberta Resende. "Qual o Impacto Do WhatsApp Em Eleições? Uma Revisão Sistemática (2010-2019)." *Revista Debates* 13 (December 16, 2019): 58–88. https://doi.org/10.22456/1982-5269.96255.

Sehat, Connie Moon. "Closed Messaging Spaces: Thinking About Best Practices and Ethics in a Global Environment." *Medium* (blog), October 4, 2018. https://misinfocon.com/closed-messaging-spaces-thinking-about-best-practices-and-ethics-in-a-global-environment-10fc82191743.

Solano Gallego, Esther. "How Bolsonaro Came to Pass." *Rosa Luxemburg Stiftung NYC* (blog), November 27, 2018. http://www.rosalux-nyc.org/how-bolsonaro-came-to-pass/.

Sumartono, Suwandi Sumartias. "Utilization of Whatsapp As A Political Communication Channel Politicians Padang City." In *Advances in Social Science, Education and Humanities Research*, 136:428–31. Atlantis Press, 2017. https://doi.org/10.2991/icosop-17.2018.67.

Telegram. "Terms of Service." Accessed July 8, 2020. https://telegram.org/tos.

Centers for Disease Control and Prevention. "The U.S. Public Health Service Syphilis Study at Tuskeegee: The Tuskeegee Timeline," 2007 2020. http://www.cdc.gov/tuskegee/timeline.htm.

Wardle, Claire, Angela Pimenta, Guilherme Conter, Nic Dias, and Pedro Burgos. "Comprova: An Evaluation of the Impact of a Collaborative Journalism Project on Brazilian Journalists and Audiences." First Draft, June 27, 2019. https://firstdraftnews.org:443/latest/comprova-an-evaluation-of-the-impact-of-a-collaborative-journalism-project-on-brazilian-journalists-and-audiences/.

We Are Social. "2018 Digital Yearbook." Internet, 17:11:28 UTC. https://www.slideshare.net/wearesocial/2018-digital-yearbook-86862930?ref=https://wearesocial.com/.

WhatsApp Blog. "Keeping WhatsApp Personal and Private." *WhatsApp Blog* (blog), April 7, 2020. https://blog.whatsapp.com/Keeping-WhatsApp-Personal-and-Private.

———. "More Changes to Forwarding." *WhatsApp Blog* (blog), January 21, 2019. https://blog.whatsapp.com/more-changes-to-forwarding.

———. "Two Billion Users -- Connecting the World Privately." *WhatsApp Blog* (blog), February 12, 2020. https://blog.whatsapp.com/two-billion-users-connecting-the-world-privately.

WhatsApp.com. "WhatsApp FAQ - How to Use Broadcast Lists." Accessed July 13, 2020. https://faq.whatsapp.com/iphone/chats/how-to-use-broadcast-lists.

WhatsApp.com. "WhatsApp Legal Info." Accessed July 8, 2020. https://www.whatsapp.com/legal/#key-updates.

Additional Resources

First Draft. *How Journalists Can Monitor TikTok to Inform Their Reporting | Webinar with Laura Garcia*, 2020. https://www.youtube.com/watch?v=n1WVwPQcwRY&list=PL0n8am2uBRCBUZI5ftm2jQ63tlYrtY4_I&index=10.

———. *Reporting on Coronavirus - Using WhatsApp to Find Covid-19 Stories | Webinar*, 2020. https://www.youtube.com/watch?v=mUfUqI8a7bs&list=PL0n8am2uBRCBUZI5ftm2jQ63tlYrtY4_I&index=2.

Melo, Philipe de Freitas, Carolina Coimbra Vieira, Kiran Garimella, Pedro O. S. Vaz de Melo, and Fabrício Benevenuto. "Can WhatsApp Counter Misinformation by Limiting Message Forwarding?" In *Proceedings of the International Conference on Complex Networks and Their Applications*. Lisbon, Portugal, 2019. http://arxiv.org/abs/1909.08740.

Reis, Julio C. S., Philipe Melo, Kiran Garimella, Jussara M. Almeida, Dean Eckles, and Fabrício Benevenuto. "A Dataset of Fact-Checked Images Shared on WhatsApp During the Brazilian and Indian Elections." *Proceedings of the International AAAI Conference on Web and Social Media* 14 (May 26, 2020): 903–8.

Resende, Gustavo, Philipe Melo, Hugo Sousa, Johnnatan Messias, Marisa Vasconcelos, Jussara Almeida, and Fabrício Benevenuto. "(Mis)Information Dissemination in WhatsApp: Gathering, Analyzing and Countermeasures." In *The World Wide Web Conference*, 818–828. WWW '19. New York, NY, USA: Association for Computing Machinery, 2019. https://doi.org/10.1145/3308558.3313688.

Appendix A: Examples of Considerations, Practices, and Questions Regarding Closed Message Research

Considerations from Sehat, 2018[72]
"We wanted to inquire about two things, which exist in apparent tension: what are the good methodologies to explore this space — in which one could expect reliable data? — And what ethics of the situation need to be taken into account, given the very privacy that grounds these spaces?"

- We need to gather more information about what sort of content gets shared and who shares it.
- Before deploying surveys or other research, think about the limits of what knowledge could be reliably gathered.
- The bar for ethical review may need to be higher than average, given the context.

Considerations and Practices from Bengani, 2019[73]:
- Don't use your personal phone or phone number.
- Identify yourself and your institutional affiliation.
- Use hard drives and not a cloud in data collection.
- Respect general privacy and have ethical cautions.

Questions from Dotto et al., 2019[74]:
- When will you disclose your identity? Will you disclose when you first enter the group, when you find something useful in the group you'd like to include in your reporting, when you have completed newsgathering in the group, or when your story is published? If you plan on being in the group for an extended period, or the group gains new members after your initial disclosure, will you re-disclose your identity?
- What is the purpose of this group? Is the group likely to be hostile, and how would group members react to a reporter within their midst? Entering a closed group that facilitates criminal activity or advocates extremist ideologies, for example, may lead to a different disclosure decision than entering a WhatsApp conversation consisting of local parents or a secret group of employees looking to unionize.
- Is your entry and presence in the group, using your real identity, likely to draw unwanted attention or abuse? Journalists of color and women, for example, may face additional security concerns when entering into certain potentially hostile groups, which may lead to a different disclosure decision.
- If you decide to enter the group using your real identity, to whom will you disclose this information? Will you disclose it to the group administrator, or to the whole group?

---

[72] Sehat, "Closed Messaging Spaces."
[73] Bengani, "India Had Its First 'WhatsApp Election.' We Have a Million Messages from It."
[74] Carlotta Dotto, Rory Smith, and Claire Wardle, "Closed Groups, Messaging Apps, and Online Ads," First Draft's Essential Guide (First Draft, November 2019), https://firstdraftnews.org/wp-content/uploads/2019/11/Messaging_Apps_Digital_AW-1.pdf?x80491.

- Will you also disclose your reasons for being in this group?
- If the group requires you to answer certain questions before admission, will you answer these questions honestly?
- Whether there are explicit confidentiality clauses in the community guidelines of the groups you are entering: How you are going to describe the methods of newsgathering in the resulting story?
- How you will go back into the group after the story's publication and share the information you have learned?
- Is there still value in news journalists dwelling in social media platforms where participants are routinely attacked before they're heard and where media manipulation is so dominant?
- And if so, can and should journalists look a little deeper into this local beat and seek greater insight from digital communities in private, invitation-only networks?
- How easy and appropriate is it to join groups and communities intent on peer-to-peer online discussion?
- What issues or barriers are there in journalists building meaningful connections in these online spaces?
- How much of what can be observed in these forums would a journalist wish to use or relay to colleagues for publication or broadcast?
- Are there ways of obtaining the information you are seeking without entering into closed online spaces?
- What do you hope to obtain from joining this group? Are you looking to find sources and tips, or to gain background knowledge to inform your reporting? Or are the existence and the content of the group itself the focus of your intended story?
- Is this a group that would expect "lurkers"? Would members reasonably expect conversations and other content from these groups to be made public?
- What is the size of each closed group you are planning on entering, and how does that affect the expectation of privacy for each group?
- Would your writing expose group members to negative consequences?
- What is the public interest in your potential story?
- Are you planning on entering multiple groups? What is the minimum number of closed spaces you can enter into to find the information you need?