The University of Texas at Austin
School of Journalism and Media
Moody College of Communication

300 W. Dean Keeton, A1000
Austin, Texas 78712-1073
512-471-1845
journalism.utexas.edu

Summary of Disinformation Research in Journalism & Media
Respectfully Submitted to the Office of the High Commissioner for Human Rights
Josephine Lukito, Ph.D & Samuel Woolley, Ph.D

[1] What do you believe are the key challenges raised by disinformation? What measures would you recommend to address them?

Disinformation production raises many human rights concerns. Though anyone can produce social media disinformation, those with substantial political and economic power have the resources to produce large-scale disinformation campaigns that blanket a country's entire communication ecology to manipulate that country's political landscape. Domestically, this discourages political opposition and violates people's freedom of opinion and expression. Internationally, the manipulation of opinions in foreign countries is a direct violation of Westphalian sovereignty. Disinformation also disproportionately targets minority populations, encouraging discrimination on the basis of race, socioeconomic status, and political opinion.

We recommend a multi-pronged approach involving three stakeholders:

- State governments and IGOs must discourage the use of state-sponsored disinformation campaigns that target domestic and foreign audiences.
- All social media platforms must take a proactive approach in removing disinformation. While some platforms (e.g., Twitter, Facebook), do try to remove disinformation on their networks, disinformation actors have migrated to smaller platforms and encrypted messaging applications (EMAs), which are less regulated (if at all).
- Journalists must develop norms to prevent disinformation from being amplified by news outlets. This includes creating protocols for verifying information found on social media.

[2a] What legislative, administrative, policy, regulatory or other measures have Governments taken to counter disinformation online and offline?

Government responses to encrypted spaces often focus on dismantling encryption. Governments should accept that disinformation is a unique problem compared to other issues on EMAs, like criminal activity or illegal content sharing, and that disinformation can be countered without dismantling user protection. This is important because EMAs are often mass adopted by peaceful activists following periods of social unrest and government crackdowns.

In some cases, governments are not taking measures to counter disinformation. Rather, they are the disinformation producers. Political parties that use disinformation are likely to employ it when governing; for example, Brazilian President Jair Bolsonaro's campaign used disinformation to get elected. Since then, he continues to amplify disinformation about COVID-19, causing direct medical harm to the Brazilian population.

The University of Texas at Austin
School of Journalism and Media
Moody College of Communication

300 W. Dean Keeton, A1000
Austin, Texas 78712-1073
512-471-1845
journalism.utexas.edu

[2b] What has been the impact of such measures on i) disinformation; ii) freedom of opinion and expression; and iii) other human rights?

The dismantling of encryption does little to combat disinformation and may actually discourage freedom of opinion and expression by eliminating a space for citizens, particularly minority groups, to communicate safely.

[2c] What measures have been taken to address any negative impact on human rights?

To our knowledge, we are not aware of government measures that target disinformation with a consideration of human rights violations and, in some cases, the governments themselves are taking active measures to manipulate their own country's political discourse, which negative impacts their citizenry's freedom of opinion and expression.

[3a] What policies, procedures or other measure have digital tech companies introduced to address the problem of disinformation?

WhatsApp has introduced a suite of self imposed regulations in the past year that changed the rules for group creation and limited message forwarding en masse.

[3b] To what extent do you find these measures to be fair, transparent and effective in protecting human rights, particularly freedom of opinion and expression?

Twitter's takedown policies seem to be arbitrary and give substantial leniency to the people with the most power on the platform. Here is an example of disinformation and bigotry manufactured by a major politician causing takedowns for everyone but the politician, including some accounts that were technically innocent of the disinformation as it was presented by the politician.

[3c] What procedures exist to address grievances and provide remedies for users, monitor the action of the companies, and how effective are they?

Some social media, including Twitter and Facebook, allow users to report messages that might be disinformation. However, users who report get little information about how their grievance has been addressed. Disinformation actors can also bypass these reporting strategies.

[4] Please share information on measures that you believe have been especially effective to protect the right to freedom of opinion and expression while addressing disinformation on social media platforms.

The University of Texas at Austin
School of Journalism and Media
Moody College of Communication

300 W. Dean Keeton, A1000
Austin, Texas 78712-1073
512-471-1845
journalism.utexas.edu

Corrective messages and labels can be effective at combating disinformation on social media platforms. However, corrective messages posted late do not help address disinformation.

[5] Please share information on measures to address disinformation that you believe have aggravated or led to human rights violations, in particular the right to freedom of opinion and expression.

WhatsApp's content-agnostic mass forwarding limits have led to no meaningful perceived benefit in any of our interviews with journalists, propagandists, or academics. In fact, the consensus appears to be that it has systematically advantaged groups that already have a propaganda infrastructure on the world's most popular and widespread messaging app.

[6] Please share any suggestions or recommendation you may have for the Special Rapporteur on how to protect and promote the right to freedom of opinion and expression while addressing disinformation.

1. Comprehensive data protections will stem the confluence of hyper targeted disinformation. For example, countries like the United States where there is a dearth of meaningful data and privacy protection have been open to industry-specific privacy regulation. Implementing protections around the use of data in politics the same way, for instance, that health data is protected will curb hyper-targeted disinformation.
2. In the United States, Peer to Peer texting spread disinformation is a rising threat to democracy that could be curbed by removing the loophole exploited by companies like RumbleUp.
3. Countries should pass strict disclosure laws around community influencers.
4. Disinformation in encrypted spaces is part of an ecosystem with other apps. Increasing access and transparency to the platform data for journalists and academics will allow the problem to be studied from the necessary multi-platform perspective.
5. We also encourage the UN to pass a resolution denouncing the use of state-sponsored troll armies either domestically (as a means of silencing the freedom of opinion) and internationally (as it is a violation of state sovereignty).

Additional Considerations: Platforms

Encrypted messaging applications (EMA): overall themes

Key Platform Findings

- None of our findings indicate a perceived change in the lived experience of those using the apps from WhatsApp's new policies
- Inability to apply content moderation to encrypted spaces

The University of Texas at Austin
School of Journalism and Media
Moody College of Communication

300 W. Dean Keeton, A1000
Austin, Texas 78712-1073
512-471-1845
journalism.utexas.edu

- Self-regulation is not a complete solution cross platform coordinated influence campaigns
- EMAs exist as a part of a multi-platform strategy, along with op eds, blogs, Facebook, Twitter, and others. This is evident in both developed and developing democracies, including India, US, Brazil, and Mexico.
- The content on this platform included coronavirus denialism/disinformation, discrediting traditional media and journalists, and higher activity during elections
- In Brazil and US, accusation of pedophilia against political opponents of ruling right-wing demagogues are spread on WhatsApp
- Content is often visual or surface level. Videos often had misleading titles.
- Notably, visual disinformation can spread more easily among illiterate app users
- Our findings suggest a manipulation of content

Findings Related to EMAs in the United States
- EMAs were a refuge for violent extremists
- EMAs facilitate disinformation in international communities
- Possible international origin of the disinformation
- Local leaders cannot identify sources of disinformation or possible coordination
- Prevalence of culturally targeted disinformation and electoral disinformation

Findings related to EMAs in Brazil
- Tax payer dollars are funneled into disinformation campaigns against enemies of current President of Brazil Jair Bolsonaro.
- WhatsApp groups automatically created from campaign data segmentation

Findings Related to EMAs in India
- Like Brazil, there are data-centric WhatsApp disinfo campaigns in India

Findings Related to EMAs in Mexico
- Paid propagandists manipulate trends and have a suite of tools for multi-platform manipulation for politicians

Findings Related to Extremists and EMAs
- As the Facebook and Twitter ecosystems take more steps to combat disinformation, extremist groups in the US have moved to EMAs, where there is no content moderation
- EMAs are used by extremist groups to push messaging onto other platforms and recruit

Political campaign apps
- One-way communication between constituents and politicians, cannot be audited by fact checkers or journalists.
- Hard to pull off except for biggest campaigns and candidates
- These campaign apps rely on an independent source of data on constituents that can be used to build lookalike audiences, etc.

The University of Texas at Austin
School of Journalism and Media
Moody College of Communication

300 W. Dean Keeton, A1000
Austin, Texas 78712-1073
512-471-1845
journalism.utexas.edu

- Organizing and comm tools for campaigns

Peer to Peer texting

- P2P is more heavily invested in in the US than EMAs
- Outdated disclosure laws + data centric campaigning + sub contractors = anonymous, mass messaging from unknown sources
- Documented cases of electoral disinformation and at least one instance of a call to action to harass the vote counting station in the Philadelphia Convention Center
- Based on a loophole in manual texting from long form numbers
- Wallet Pass in development to stay ahead of regulation

Additional Considerations: Key Actors

Influencers
- Influencers are the new community organizers, especially as social media platforms like instagram have become more political in the past year
- Small-scale influencers (<10k followers), in particular, have become increasingly popular amongst both commercial brands and political campaigns and interest groups. This is due to their increased engagement with followers, increased trust, cheaper prices and targetable, niche audiences.
- Influencers coordinate amongst themselves through engagement pods and collectives, like hype houses
- Influencers are influenced by their followers, shaping norms around political movements, covid information, and conspiracies like QAnon

State-Sponsored Disinformation Actors
- State-sponsored disinformation campaigns coordinate messages across multiple platforms, including social media platforms, websites, and blogs.
- To combat this disinformation flow, it is necessary to create tools that can tag and block disinformation messages as they pass from one platform to the other.

Journalists
- Disinformation actors target journalists and news organizations to amplify disinformation messages and to create distrust in a country's media ecology.
- Disinformation actors impersonate journalists or newsrooms on social media.
- Disinformation actors try to communicate with genuine news organizations to get their messages embedded in news content. In the United States, we found that a majority of news organizations accidentally quoted a Russian troll in at least one news story.

The University of Texas at Austin
School of Journalism and Media
Moody College of Communication

300 W. Dean Keeton, A1000
Austin, Texas 78712-1073
512-471-1845
journalism.utexas.edu

- When a message from a disinformation actor is embedded in a news story, it raises the perceived trustability of the disinformation actor. As a result, the disinformation actor will get more followers on social media.

Political data brokers
- Politics is a testing ground for proof of concept with targeting, but firms move quickly to commercial contracts for the big money
- Data brokers admit to being uncomfortable with the level of information
- The real insights come from layering many types of data and creating models, but requires money that not all campaigns have
- Move towards consolidating data troves along party lines to strengthen campaigns

Additional Considerations: Misc.

Geopropaganda: general
- Location data is very commonly used, but most often in combination with other data like voter files or demographic data
- Location data can be used to geo-fence a particular region and target all who are in that region or who have visited that region
- Geofencing rallies and combining with voter data, as Brad Parscale did, to see who has previously voted, and for what party
- Geofencing homes of donors, lobbyists and legislators and geofencing capitol buildings
- Costs money to do it right so this isn't the highest priority of all campaigns

Data privacy
- States need to pass a federal privacy law that is human-centered and protects citizens with or without commercial interaction.

Protest Surveillance
- The level of response by law enforcement (whether at the local, state, and federal level) can vary according to the demographic and ideology of those participating in the protest, demonstration, march, etc. Thus, the notion that surveillance creates "accountability" is wholly, if not at least in part, contingent upon social and political factors that are exacerbated by surveillance itself.

Cross-Platform Coordination
- Information that flows from one platform to another can "transform," making it harder to identify. For example, text on one social media platform may be saved as a computer screenshot (a saved image of a computer screen) and shared on another platform as an image.

The University of Texas at Austin
School of Journalism and Media
Moody College of Communication

300 W. Dean Keeton, A1000
Austin, Texas 78712-1073
512-471-1845
journalism.utexas.edu

- The easiest cross-platform disinformation flows that can be studied or removed are hyperlinks and URLs that are tied to disinformation websites. By identifying social media accounts that share URLs from disinformation websites often, industry professionals and academics may be able to better identify disinformation campaigns.

Links to Our Team's Research

1. Encrypted Messaging Apps: https://mediaengagement.org/research/encrypted-propaganda/
2. Location Targeting: https://www.brookings.edu/techstream/political-operatives-are-targeting-propaganda-by-location/
3. Voter Surveillance: https://www.technologyreview.com/2020/06/21/1004228/trumps-data-hungry-invasive-app-is-a-voter-surveillance-tool-of-extraordinary-scope/
4. Journalism and Disinformation: https://journals.sagepub.com/doi/abs/10.1177/1940161219895215
   a. Impact of disinformation tweets in US news stories: https://academic.oup.com/joc/advance-article-abstract/doi/10.1093/joc/jqaa042/6104044
5. Peer-to-peer texting: https://mediaengagement.org/research/peer-to-peer-texting-and-the-2020-election/
6. Social media influencers: https://mediaengagement.org/research/social-media-influencers-and-the-2020-election/
7. Identity Manipulation by Disinformation Actors: https://www.tandfonline.com/doi/abs/10.1080/1369118X.2019.1621921