

The Mitigation of Misinformation and the Preservation of Human Rights

In 1989, Sir Tim-Berners Lee, a British computer scientist, invented what we know today as the World Wide Web. And with the birth of our most cherished gift from the digital revolution came an assortment of new dangers and responsibilities. Today, the internet and social media serve as a petri dish in which misinformation grows and spreads. This type of information is not just misleading, it is also severely damaging to democracy. Disinformation, a form of misinformation which is deliberately deceptive, is even more harmful. In the United States, this “infodemic” has created misinformed voters and fostered extreme party polarization in a country that is already more divided than ever. In the wake of the most turbulent year for our political and public health systems, social media companies are struggling to filter inaccurate and outright deceptive content from their platforms. This is a huge undertaking; this type of content is not only hard to identify, but users’ rights can easily become caught in the crossfire when battling it. Companies sometimes struggle to mitigate misinformation without hindering free speech. To tackle the widespread misinformation on social media while preserving this freedom, companies should develop both preventative and reactive policies while continuously checking for possible human rights infringements.

To fully understand the threat that misinformation poses to society, one must explore its past. Misinformation has a long history of creating chaos and disrupting century-old institutions as well as just tricking people into believing false rumors. In an article about the role of disinformation in elections, Max Bader from Leiden University addresses the problem as it has appeared in elections past. He discusses both domestic disinformation as well as disinformation spread on social media by foreign culprits. Bader identifies the 2016 Trump campaign and the Brexit campaign as two of the most disinformation-ridden campaigns in history. There is evidence that the Trump campaign involved manipulation of social media coverage, and that

social media bots influenced opinions in the Brexit campaign. The 2016 Trump campaign resulted in a presidency widely agreed to be one of the most damaging in American history, and misinformation is likely to blame for that election outcome. A smaller, less serious example from the beginning of the coronavirus outbreak demonstrates the wildfire-like expansion of misinformation. In March of 2020, several videos of dolphins swimming in a Venetian canal were shared along with messages of hope for a struggling Italy. Dolphins hadn't been spotted in those canals for decades, supposedly because of boat pollution which had cleared up due to the shutdowns. The story brought joy to thousands of people across the globe. Except there were no dolphins in Venice; the videos were actually shot hundreds of miles away in a Mediterranean port. While this fake story was not harmful to those who believed it, it demonstrates how easily misinformation (including the harmful kind) can gain traction and become widely accepted as truth.

Retrospectively, we wonder how these instances of misinformation could have been stopped before they became so widespread. The solution to this problem is complex, but it begins with social media companies implementing platform policies that protect their users from misleading information. It is impossible to identify all posts spreading falsehoods, but companies can implement techniques that help cut down on them. One technique, called fact-checking, is commonly seen on platforms like Facebook and TikTok. Fact-checking is reactive; it involves flagging posts that include keywords like "Trump," "coronavirus," or "hoax" with a message alerting the viewer that the post is potentially misinformed. The flag is intended to make viewers think twice before believing the statement or sharing it with their friends. In some cases, a link is provided to a reliable source like the CDC. This is an effective way to slow the dissemination of false information without silencing users. A second, less well-known method of fighting misinformation actually involves prohibiting a practice rather than implementing one. The practice is known as microtargeting. According to former Chair of the U.S. Federal Election Commission Ellen Weintraub, microtargeting "amplifies disinformation harms by targeting

susceptible groups with political ads that the rest of society does not see.” Microtargeting not only limits certain groups’ exposure to opposing and varied ideas, but it shields entities who make untrue claims from the type of backlash they would receive from a wider, more diverse audience. Some social media companies have chosen to limit or entirely ban microtargeting, making a proactive effort to stop inaccuracies from spreading. Like fact-checking, banning microtargeting does not hinder users’ free speech. The entity making the post is not silenced. In fact, they are given a wider audience- just one that is less easily manipulated. Misinformation is not eliminated directly using this method, but it can be eliminated indirectly by creating higher consequences for those who post false information.

Regardless of the approach companies take to detecting and eliminating misinformation on their platforms, they are responsible for continuously checking their policies to ensure they do not infringe upon users’ rights. Standards such as those outlined in the United Nations Guiding Principles on Business and Human Rights (UNGPs) are helpful to social media companies making big policy changes. In her article “In a World of ‘Fake News,’” Evelyn Mary Aswad says that UNGPs “reflect international expectations for how companies should act when their operations intersect with human rights issues.” A large portion of the guide is dedicated specifically to addressing how companies should handle misinformation without obstructing the right to free speech. Several of the most popular social media platforms, including Facebook, have modeled their policies after the UNGPs. While curbing the presence of misinformation on social media is vital to democracy and societal prosperity, preserving basic human rights is of paramount importance. With careful consideration and a strong dedication to protecting user rights, social media companies can develop effective and fair policies that achieve both goals.

Bibliography

- Aswad, Evelyn Mary. "In a World of 'Fake News,' What's a Social Media Company to Do?" *Utah Law Review*, no. 4, 2020, pp. 1009-10028. *EBSCOhost*, <http://ezproxy.clark.edu:12177/ehost/pdfviewer/pdfviewer?vid=6&sid=6af2fdcc-950a-4ea8-9b40-c9af9bd79afb%40pdc-v-sessmgr02>. Accessed 29 March 2021.
- Bader, Max. "Disinformation in Elections." *Security and Human Rights*, vol. 29, 2018, pp. 24-35. *EBSCOhost*, https://brill.com/view/journals/shrs/29/1-4/article-p24_24.xml.
- Ciampaglia, Giovanni Luca, and Filippo Menczer. "Biases Make People Vulnerable to Misinformation Spread by Social Media." *Scientific American*, 21 June 2018, <https://www.scientificamerican.com/article/biases-make-people-vulnerable-to-misinformation-spread-by-social-media/>.
- Daly, Natasha. "Fake animal news abounds on social media as coronavirus upends life." *National Geographic*, 20 March 2020, <https://www.nationalgeographic.com/animals/article/coronavirus-pandemic-fake-animal-viral-social-media-posts>.
- "History of the Web." World Wide Web Foundation, <https://webfoundation.org/about/vision/history-of-the-web/>.
- Ray, Siladitya. "TikTok Will Flag And Limit Spread Of Videos In Case Of 'Inconclusive' Fact Checks." *Forbes*, 3 February 2021, <https://www.forbes.com/sites/siladityaray/2021/02/03/tiktok-will-flag-and-limit-spread-of-videos-in-case-of-inconclusive-fact-checks/?sh=1aac0323796e>.
- Ruggie, John G. "FACEBOOK IN THE REST OF THE WORLD." *Harvard Kennedy School*, 2018, pp. 1-2, https://media.business-humanrights.org/media/documents/files/documents/John_Ruggie_Facebook_15_Nov_2018.pdf.

Vraga, Emily K., and Leticia Bode. "Correction as a Solution for Health Misinformation on Social Media." *American Journal of Public Health*, vol. 110, 2020, pp. S278-S280. *EBSCOhost*, <http://ezproxy.clark.edu:12177/ehost/detail/detail?vid=8&sid=6af2fdcc-950a-4ea8-9b40-c9af9bd79afb%40pdc-v-sessmgr02&bdata=JnNpdGU9ZWwhvc3QtbGl2ZSZzY29wZT1zaXRl#AN=146194632&db=a9h>.