About mis-disinformation, its potential impacts, and the challenges to finding effective countermeasures

NICOLAS RUTHERFORD, ANALYST | FEBRUARY 2023

Laboratoire sur l'intégrité de l'information

Information Integrity Lab

uOttawa

Disinformation, Misinformation and Fake News

In contemporary discourse, terms like "disinformation" and "misinformation" have become catch-all terms for a variety of things, such as news that is hyperpartisan, clickbait content, rumours, conspiracy theories, and even ideological biases. Because of this, "the sense prevails that no two people who talk about disinformation or misinformation are talking about quite the same thing."[1] To gain a clearer understanding of the terms, researchers like Wardle (2018) propose some useful definitions[2] :

First, "disinformation" refers to verifiably false information that is created and spread intentionally in order to confuse, manipulate or mislead. It may contain a blend of truth and untruth, or purposely exclude/distort context, making it harder to differentiate from content that is "authentic". Examples of disinformation include the politically motivated claims that voter fraud influenced the outcome of the 2020 US Presidential Election, despite numerous independent audits and court cases finding no tangible evidence that supports this[3], or the narrative that the COVID-19 virus was a "hoax" orchestrated by global institutions and governments in collusion with Big Pharma.[4]

IN RECENT YEARS, TERMS LIKE "FAKE NEWS" HAVE ALSO BECOME INCREASINGLY POPULAR.

"Misinformation" on the other hand designates incorrect or misleading information that is not intentionally deceptive and which hasn't been shared in order to cause harm. A common example is when, in the midst of a breaking news event that's unfolding and being talked about on social media, "people share rumours or old photos, not realizing that they're in fact not connected to [those] events."[5]

In recent years, terms like "fake news" have also become increasingly popular.[6] The expression refers to fabricated stories that have been deliberately designed to imitate real news articles, using similar presentation styles and formats. This makes them more convincing to readers and viewers, but also more challenging to identify as false in the first place. As usage of the phrase has grown, so too have criticisms towards it, since it's often misused to discredit opposing viewpoints and to undermine the credibility of professional news media around the world.[7] It's also not very descriptive of most of the mis-disinformation that one encounters, especially online, because in reality "most of [this] does not even masquerade as news. It is memes, videos, images or coordinated activity on Twitter, YouTube, Facebook or Instagram. And most of it isn't fake; it's misleading or, more frequently, genuine, but used out of context."[8]

Researchers like Tworek (2021) point out that the production and spread of mis-disinformation isn't at all a new phenomenon, and that in fact it's been an integral feature of the media environment for news, for a very long time.[9] Complaints about "faking in news", as well as pointing the finger at "reader gullibility" and "sensationalist profit", can be traced back to at least the start of newspapers themselves in the seventeenth century.[10] Comparable challenges existing in the present-day suggests that these issues are in reality persistent and did not arise solely due to the emergence of the internet or other modern information technologies. That being said, the ongoing digital transformation of our society and methods of communication has caused notable shifts which are new and require our attention ; one such issue is the fact that today, mis-disinformation content that appears credible can not only be produced by virtually anyone, it can also be disseminated at breakneck speeds and on a scale that is truly global.

Mis-disinformation in the 21st century "information environment"

The current information and technology landscape seems to be a more favourable space for mis-disinformation to exist and thrive in than the one we had even just a decade ago. Among other things, it encourages extreme levels of information consumption - accessed most often through devices like smartphones - leading to a near constant personal "news cycle". At the same time, the democratisation of information providers has allowed many new voices - not all of them intent on fact-checking their claims - to gain platforms and a potential access to very large audiences. Information environments are also being re-territorialized by the algorithm-driven emergence of informational echo chambers, which is limiting the cross-pollination of ideas and perspectives and, in recent years, has noticeably heightened social and political tensions.

Within this landscape, mis-disinformation holds a significant potential to mislead any person who comes across it, though it should be noted that the precise impact that manipulated information can have on a person in terms of changing their mind, has been hard to measure and fully explain.[11] What is clear is that in an information environment that is chaotic and overwhelming - and where mis-disinformation proliferates in the absence of adequate checks or barriers - it can become especially difficult for people to distinguish truth from falsehood. Such circumstances also seem to correlate with a variety of "bad social outcomes" emerging, like :

- An erosion of trust in government institutions, scientific research, and mainstream journalism media, driving some members of the public away from content or sources which they don't perceive are in alignment with their own ideological perspectives (intensifying echo chamber effects), while sometimes also driving them to simply consume less information overall and even to sever social relationships.[12]

- The undermining of significant social and political functions, including collective decision-making processes in democratic societies, (1) by skewing public discussions away from important issues and topics (e.g. What immediate action should we take on climate change?), (2) by casting doubt on the legitimacy of electoral processes and their outcomes, (3) and in some instances, even by disseminating targeted disinformation to dissuade voter turnout among particular demographics (e.g. In the USA, false voting information being sent through text messaging to residents of several predominantly African-American communities in Alabama during the state's Senate special election in 2017).[13]

- The promotion and amplification of harmful conspiracy theories and hate speech, and by extension of the fringe/extremist groups that relay them, leading in certain cases to acts of physical violence (e.g. PizzaGate, the January 6th attack on the U.S. Capitol, etc.). In situations of conflict or war, mis-disinformation has also been used to dehumanize the "enemy". For example, the portrayal of the Rohingya people as illegal immigrants and terrorists by the Myanmar government created the false and inflammatory perception that they were a threat to the nation and thus justified targets for the violent attacks - and genocide - perpetrated against them in 2017.[14]

About the challenges related to deploying effective countermeasures to mis-disinformation

Today, governments, technology companies and civil society groups around the world recognize that mis-disinformation is a major problem, and yet all of them are still struggling to find and implement effective responses. There can be many reasons for this, like the fact that it's become easier for malign actors – who range from individual trolls to (domestic or foreign) governments – to produce disinformation content that is sufficiently convincing. Indeed, wider access to inexpensive or even free media manipulation tools has allowed more people than ever to create and share mis-disinformation quickly, cheaply, and more or less efficiently, especially in the online space. On the popular social media app Tik Tok, for instance, nearly 20% of the videos made and shared by users have been found to contain mis-disinformation.[15]

Another aspect relates to the particular medium through which mis-disinformation content is generated and spread. In the past, a greater proportion of the false and misleading information that people came across was presented in textual form, whereas nowadays, it is more likely to incorporate multiple modes of communication, such as text, images, speech, and video - making it "multimodal". Mis-disinformation in this form is not only easier to consume, but research indicates that it may also have a more significant impact than its purely text-based counterpart because audiovisual media in particular is "more attention-grabbing and emotionally engaging than textual information."[16] It can also be more complicated to detect, since often this type of mis-disinformation content involves the repurposing, or recycling, of authentic images and speech, which are then presented in a misleading way rather than directly manipulated or fabricated from scratch. Automated moderation tools, which are essential for dealing with online mis-disinformation content at very large scales, thus face a significant obstacle in terms of accurately differentiating between repurposed or out-of-context content and its original, authentic sources. An additional issue is how to determine which of these materials constitute mis-disinformation instead of legitimate satire, art, or other valid forms of expression.
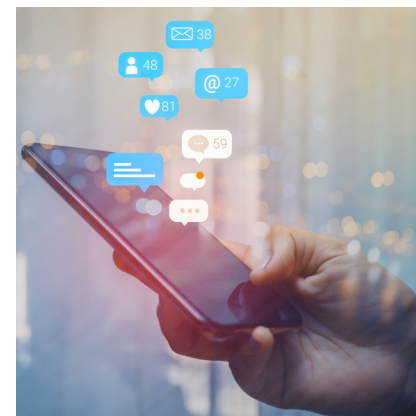
Meanwhile, recent innovations in the budding field of generative AI are making it possible for malign actors to create entirely new image and video content that is highly believable ; with the right AI-enabled tools, anyone can replace faces and speech in video to make it appear as if someone said or did something that never happened.[17] The arrival onto the scene of tools such as Open AI's ChatGPT and DALL-E which, respectively, are able to generate text and image content in response to command prompts from users, as well as Microsoft's voice-mimicking tool VALL-E, make it even easier to produce highly customized and hyper-realistic synthetic media in a matter of seconds, opening the door further to all kinds of nefarious uses. If current and past trends are anything to go by, bad actors engaging in computational propaganda activities will be able to count on newer and better tools becoming available

NEARLY 20% OF THE VIDEOS MADE AND SHARED BY USERS HAVE BEEN FOUND TO CONTAIN MISDISINFORMATION.

for misuse in the future. This puts governments, technology companies and civil society groups all over the world in a position where they must continuously reassess and adjust their response strategies. In this escalating "arms race" over protecting the integrity of information, it is difficult for those who are attempting to mitigate the problem to durably gain the upper hand.

One final and crucial factor that makes mis-disinformation very challenging to address is its psycho-social dimension. Part of the reason it can be so effective is because our current information environment enables the easy exploitation, reward, and amplification of several cognitive biases that are widespread among the general public. One such bias is our inclination to gravitate towards, and be more likely to disseminate, news and other information that aligns with our preexisting (mis)beliefs. As a result, information that is factually incorrect or misleading, but that "feels right", can be validated as legitimate and truthful through an emotional resonance effect.[18] As Wardle and Derakhshan (2017) explain, "when most social platforms are engineered for people to publicly 'perform' through likes, comments or shares, it's easy to understand why emotional content travels so quickly and widely, even as we see an explosion in fact-checking and debunking organizations."[19]

Looking ahead

Although progress is being made when it comes to detecting and protecting against instances of "information pollution", it should be clear that eradicating all of the threats posed by mis-disinformation is not a realistic expectation. "Solving" the issue is not one simple "mis-disinformation law" away, and no "one size fits all" approach is likely to work, especially in the face of mis-disinformation phenomena that are widespread, complex and continually evolving. The reality is that governments, technology companies, civil society groups and all other stakeholders must keep finding ways to adapt and better cope with mis-disinformation, because it is here to stay.[20]

One way in which various parties are trying to confront the issue is to break it down into smaller parts and work on, for instance, ways in which we can mitigate the amplification of poor-quality news in our media environments, or how to empower individuals to become more critical and resistant to fake news. Innovations can also be made to enhance the visibility and spread of quality information in our information ecosystems in the first place.

Various specific "solutions" are already being implemented, such as creating online community fact-checking tools, introducing social media policies that eliminate the economic incentives associated with sharing disinformation, promoting culturally adapted media literacy initiatives, lobbying governments to pass laws or regulations targeting those who disseminate false information, or developing AI-automated systems that can more accurately detect and remove problematic content. Each of these is meant to tackle either the "supply side" of mis-disinformation, meaning the sources of false or misleading information, or the "demand side" which is concerned with those who consume it. None of the measures or initiatives mentioned above are sufficient on their own, and some, particularly those that pertain to regulating speech, can be highly controversial.[21]

Ultimately, addressing the "mis-disinformation problem" should be seen as an ongoing process, and one which requires additional research, collaboration and innovation across multiple disciplines, including, but not limited to, computer and information science, psychology, journalism, political science, law, anthropology, etc. Actors in government, the technology sector and civil society should be working together in order to find new ways to promote and protect information integrity as well as to undermine those environmental factors that are facilitating the spread of mis-disinformation in our information ecosystems.

GOVERNMENTS, TECHNOLOGY COMPANIES, CIVIL SOCIETY GROUPS AND ALL OTHER STAKEHOLDERS MUST KEEP FINDING WAYS TO ADAPT AND BETTER COPE WITH MIS-DISINFORMATION, BECAUSE IT IS HERE TO STAY.

Endnotes

1. Bernstein, J. (2021, September). Bad News. Harper's Magazine. URL : https://harpers.org/archive/2021/09/bad-news-selling-the-story-of-disinformation/

2. Wardle, C. (2018). Information disorder: The essential glossary. Harvard, MA: Shorenstein Center on Media, Politics, and Public Policy, Harvard Kennedy School. URL : https://firstdraftnews.org/wp-content/uploads/2018/07/infoDisorder_glossary.pdf

3. Vasilogambros, M. (2020). Election Disinformation Fears Came True For State Officials. Pew. URL : https://www.pewtrusts.org/en/research-and-analysis/blogs/stateline/2020/11/20/election-disinformation-fears-came-true-for-state-officials

4. Fact check: The COVID-19 pandemic was not orchestrated by pharmaceutical companies, investment groups and philanthropists. (2021, January 30th). Reuters, Everything news.

5. Wardle, C. (2020). The Age of Information Disorder. In Craig Silverman (Ed.), Verification Handbook : For Disinformation and Media Manipulation (p. 9). European Journalism Center. URL : https://s3.eu-central-1.amazonaws.com/datajournalismcom/handbooks/Verification-Handbook-3.pdf

6. "Fake news" was declared the "word of the year" by dictionary publisher Collins in 2017.

7. Habgood-Coute, J. (2018). The term "Fake news" is doing great harm. The Conversation. URL : https://theconversation.com/the-term-fake-news-is-doing-great-harm-100406

8. (Wardle, 2020, p. 7)

9. Tworek, H. (2021). Disinformation : It's History. Centre for International Governance Innovation. URL : https://www.cigionline.org/articles/disinformation-its-history/

10. Ibid.

11. Tucker, J. A., Guess, A., Barbera, P., Vaccari, C., Siegel, A., Sanovich, S., Stukal, D., & Nyhan, B. (2018). Social media, political polarization, and political disinformation: A review of the scientific literature. URL : https://ssrn.com/abstract=3144139

12. Colomina, C., Margalef, H. S., Youngs, R., & Jones, K. (2021). The impact of disinformation on democratic processes and human rights in the world. Brussels: European Parliament. URL : https://www.europarl.europa.eu/RegData/etudes/STUD/2021/653635/EXPO_STU(2021)653635_EN.pdf

13. Popken, B. (2020, September 2). Digital vote suppression efforts are targeting marginalized groups, report warns. NBC News. URL : https://www.nbcnews.com/tech/security/digital-vote-suppression-efforts-are-targeting-marginalized-groups-report-warns-n1239133

14. Amnesty International. (2022, September 29). Myanmar : Facebook's systems promoted violence against Rohingya ; Meta owes reparations. URL : https://www.amnesty.org/en/latest/news/2022/09/myanmar-facebooks-systems-promoted-violence-against-rohingya-meta-owes-reparations-new-report/

15. Brewster, J., Arvanitis, L., Pavilonis, V., Wang, M. (2020, September). Beware the 'New Google:' TikTok's Search Engine Pumps Toxic Misinformation To Its Young Users. Newsguard. URL : https://www.newsguardtech.com/misinformation-monitor/september-2022/

16. Powell T. E., Boomgaarden H. G., De Swert K., de Vreese C. H. (2015). A clearer picture: The contribution of visuals and text to framing effects. Journal of Communication, 65, 997–1017. URL : https://doi.org/10.1111/jcom.12184

17. One notorious example is Jordan Peele's Obama deepfake, which went viral in 2018.

18. Ecker, U.K.H., Lewandowsky, S., Cook, J. et al. (2022). The psychological drivers of misinformation belief and its resistance to correction. Nat Rev Psychol 1. pp. 13–29. URL : https://doi.org/10.1038/s44159-021-00006-y

19. Wardle, C., & Derakhshan, H. (2017). Information disorder: Toward an interdisciplinary framework for research and policymaking (Vol. 27, pp. 1-107). Strasbourg: Council of Europe. p.7. URL : https://rm.coe.int/information-disorder-report-version-august-2018/16808c9c77

20. Van Raemdonck, N., & Meyer, T. (2022). Why Disinformation is Here to Stay. A Socio-technical Analysis of Disinformation as a Hybrid Threat. In L. Lonardo (Ed.), Addressing Hybrid Threats: European Law and Policies Edward Elgar.

21. Nielsen, R. K. (2020, December 17). How to respond to disinformation while protecting free speech. Reuters Institute for the Study of Journalism. URL : https://reutersinstitute.politics.ox.ac.uk/news/how-respond-disinformation-while-protecting-free-

Laboratoire sur l'intégrité
de l'information

Information Integrity Lab

uOttawa