COMMUNIQUÉ

Canadian Society for the History and Philosophy of Science

Société Canadienne d'Historie et de Philosophie des Sciences

Nº 105 autumn/l'autome 2022

Media

Wrong on the Internet: Why some common prescriptions for addressing the spread of misinformation online don't work

Isaac Record & Boaz Miller

Leading prescriptions for addressing the spread of fake news, misinformation, and other forms of epistemically toxic content online target either the platform or platform users as a single site for intervention. Neither approach attends to the intense feedback between people, posts, and platforms. Elsewhere, we provide an account for what we believe is a more productive approach (Record and Miller, 2022). Here we will specify what goes wrong in most approaches, an exercise that is worthy in itself.

Platform-centered approaches note a duty of platforms, such as Facebook and Twitter, to act as responsible gatekeepers, to monitor, and to filter misinformation (e.g., Gillespie 2018). In their consideration of platform duties, O'Conor & Weatherall write:

Algorithmic responses can help, but more is needed: ultimately, we need human editorial discretion, armies of fact checkers, and ideally, full financial and political independence between the groups whose actions are covered by news organizations, whose platforms are used to distribute news broadly, and who are responsible for evaluating whether claims are true (2019, 184).

Some hope that the state will coerce platforms to fulfill this duty. O'Conor & Weatherall add that "part of the picture will have to involve regulatory bodies in government as well as online sources whose entire purpose is to identify and block sources of misinformation" (2019, 184).

We have four objections to this approach. First, expecting corporations to adopt responsible practices or hoping for salvation by the state is the stuff of dystopian fiction. Giving monopolistic mega-corporations the license to filter content for truth is a cure worse than the disease. AndMoreover, states do not have a good track record regulating free flow of politically inconvenient information, and they may be even less trustworthy than private corporations (Origgi 2013; Tufekci 2017).

Second, while platforms are already expected to filter incitements to violence and pornography, and we may arguably also expect them to filter institutionally organized attempts to spread misinformation and fake sites that impersonate legitimate sites, it is less clear that we should also expect them to epistemically monitor posts that individuals make from their private accounts. Such posts widely vary in their level of factual accuracy, and many are in an epistemic grey area. Monitoring epistemically toxic content requires extensive inquiry and subtle, contextual judgment, which platforms appear incapable of doing, bearing in mind their abysmal track record at transparently monitoring posts for offensive content or in giving users proper channels to appeal their decisions (Vaccaro et al. 2020; Schwarz 2019).

Third, making platforms epistemic gatekeepers is an attempt to "return" to an imagined past-century media environment that never quite existed, in which editors and curators alone decided what was news. We share a concern about divided attention and a lack of common ground, both of which impede democratic decision making, but we do not think algorithmic silencing of dissenting voices can produce legitimate consensus.

Finally, such 'magic bullet' thinking misunderstands the autonomy of platform users to interpret and engage with posts and platforms. Thinking that if only we had a magic 'truth' button, no one would tell lies on the Internet anymore "prioritizes causal effects on user activity while disregarding the structural influence of problematic patterns in media messaging and representation" (Marwick 2018, 485). In fact, however, as we stressed, users have numerous strategies for reading and sharing content to weave it into their preferred narrative.

footnotes

Another common approach to misinformation is to pillory individual users for sharing stories that critics find problematic. As of this writing, media scholars have traced much bizarre content-spreading activity to QAnon and the so-called 'Pizzagate' fiasco, in which a wacky theory that Hilary Clinton was running a child slavery ring out of a pizza parlour gained circulation and was well-enough believed that an armed man assaulted the location hoping to free the children. It is sometimes suggested that this problem could be solved if individuals clicked through, fact-checked, verified sources, or otherwise employed the methods of media literacy (cf. Priest 2014).

Calls for media literacy assume that when a problematic post is re-shared, the re-sharers have made a mistake – these gullible saps have been taken in by tricksters and if they just knew a bit of critical thinking they would not make those mistakes. This does happen, but individual ignorance, generated by simple not knowing, is not the best explanation for the widespread circulation of misinformation in this context. Social media audiences are far from media illiterate. People simply aren't always so concerned about truth. Some Ppeople aren't looking on social media aren't just looking for what is true. They are also looking to share their identity and innumerable other things (Marwick 2018; Introne et al. 2018). Thus, narrow prescriptions like fact-checking will be ineffective because they assume people would do otherwise if they simply "knew better."

Fact-checking-based solutions underestimate the autonomy of audiences and the diversity of values that enter into social media engagements. "Verrit, […] Snopes, Politifact, and a host of other fact-checking sites, reflect fundamental misunderstandings about how information circulates online, what function political information plays in social contexts, and how and why people change political opinions" (Marwick 2018, 475). In particular, fact-checking sites assume that the audience for a post cares about its veracity when they often care only about verisimilitude. Posts are often polysemic, and a skilled audience members can bend nearly any messages to fit their purposes. The audience may care much more about identifying themselves with the poster through mimicking affect or simply clicking 'like' than about truth.

There is another problem with fact checking. Epistemically toxic content encompasses more than false information. It also consists of misleading information, including true information framed misleadingly or blown out of proportion. A viral story – true or false – can swamp other news. Politicians know that it is easy to distract from complicated bad news by drumming up conversation about minor but easily grasped events. A focus on fact-checking lets the liars set the agenda for our attention.

Yet another challenge is that following the routines of media literacy is burdensome. Besides liking and scrolling, sharing is the easiest operation to carry out on social media platforms. It's much easier than reading or clicking through. Requiring people to take on a burdensome task that may be orthogonal to their purpose in engaging with a post (see above) is not likely to work (Miller and Record 2017). An exception is when a person takes on a role, e.g., as a journalist or area expert, where their role responsibilities require them to carry out this task regardless of difficulty. Likewise, individuals sometimes take on the mission of fighting misinformation on a specific topic of personal interest. Notwithstanding, it is hard to see society adopting fact-check oriented media literacy practices as a general duty, especially when platform-provisioned operations such as 'like' are much easier than the alternative.

Another proposed solution is ranking posts or people. Linking present treatment to past behaviour adds a consequence for bad behaviour online. Successful examples of ranking include some Reddit communities. These typically exist inside discrete contexts where there are clear community norms. It is unclear how such a system could work in a single, undifferentiated context like Facebook or Twitter. Even with clear criteria, applying rules to ambiguous content is hard. Poe's "Law," which states that it is impossible to

to distinguish between a true believer and satirist, points to the difficulties of judging intent or affect on the Internet is difficult, and, thus, so too is establishing blame for a bad outcome. This challenge is amplified when there is no stable media ideology to help adjudicate disagreements. People's reliability also varies between subject areas, so an overall ranking based on an individual social media user may not be appropriate. Finally, people could exploit ranking and flagging mechanisms to bring down posts they don't like, such as political posts with which they disagree.

Leading prescriptions boil down to the suggestion that we make social media more like traditional media, whether by making platforms take active roles as gatekeepers, or by exhorting individuals to behave more like media professionals. Both approaches are impracticable, but there are two further reasons to resist them. First, such reforms give up the defining features of social media, foremost, its accessibility. Social media provides means for marginal voices to find audiences. Gatekeepers tend to suppress marginal voices. Second, traditional media has not escaped the ills of epistemically toxic content, so making social media look more like it isn't guaranteed to fix the problems. According to recent research (Benkler et al. 2020; Allen et al. 2020), traditional news organizations are responsible for the majority of discussion and circulation of misinformation. This is partly due to their overzealous pursuit of the very media literacy prescriptions described above, and partly due to their desperate engagement with the same attention economy as social media. Discussing crackpot theories on the nightly news, even for debunking them, makes the fringe mainstream.

Malicious misinformants exploit two weaknesses in journalism. One is the weakening business model and shortening news cycle. There isn't time, budget, or demand for in-depth reporting. Anticipating or 'prebunking' a coming flood of falsehoods is hard. Second is the fundamental tenet of journalism: to cover the story. Journalists can choose how to cover the story, e.g.: fact-check it or feature 'both sides' (which itself can be problematic when only one is legitimate, as in certain matters of expertise) (Boykoff 2007). But once a story is in circulation, news organizations feel obliged to cover it. Misinformants count on this; they aim to create viral popularity that launches them into mainstream news coverage and thereby shifts the 'Overton window' on what is appropriate and important to discuss. Interested viewers then follow the trail back down the rabbithole, often carefully documented by journalists trying to debunk the claims. In an era of information scarcity, coverage is a virtue. In an era of attention scarcity, coverage becomes a vice.

We have briefly explored some challenges of addressing epistemically toxic content in social media, and have argued that leading prescriptions, focusing on platforms or people as isolated sites for intervention, will not work. We would like to suggest that a better approach would be to make changes to both the norms governing individual posters and the platform-provisioned operations. Ideally, these changes could work in concert to address the spread of toxic information online. Elsewhere, we expand on this suggestion (Record and Miller, 2022).

References

Allen, Jennifer, Howland, Baird, Mobius, Markus, Rothschild, David, and Watts, Duncan J. (2020). 'Evaluating the Fake News Problem at the Scale of the Information Ecosystem'. *Science Advances*, advances.sciencemag.org/content/6/14/eaay3539

Benkler, Yochai, Tilton, Casey, Etling, Bruce, Roberts, Hall, Clark, Justin, Faris, Kaiser, Robert Jonas, and Schmitt, Carolyn. (2020). 'Mail-In Voter Fraud: Anatomy of a Disinformation Campaign'. *Berkman Center Research Publication No. 2020-6*, (October 2), dx.doi.org/10.2139/ssrn.3703701

References

Allen, Jennifer, Howland, Baird, Mobius, Markus, Rothschild, David, and Watts, Duncan J. (2020). 'Evaluating the Fake News Problem at the Scale of the Information Ecosystem'. *Science Advances*, advances.sciencemag.org/content/6/14/eaay3539

Benkler, Yochai, Tilton, Casey, Etling, Bruce, Roberts, Hall, Clark, Justin, Faris, Kaiser, Robert Jonas, and Schmitt, Carolyn. (2020). 'Mail-In Voter Fraud: Anatomy of a Disinformation Campaign'.

Berkman Center Research Publication No. 2020-6, (October 2), dx.doi.org/10.2139/ssrn.3703701

Boykoff, Maxwell T. (2007). "Balance as bias: global warming and the US prestige press." *Global Environmental Change* 14(2): 125-136, doi.org/10.1016/j.gloenvcha.2003.10.001

Frost-Arnold, Karen. (2021a). "Epistemic Trash and Toxic Content". Invited virtual keynote at Beyond Fake News: Mitigating the Spread of Epistemically Toxic Content Workshop by Zefat Academic College and University of Haifa, April 2021.

—— (2021b). "The Epistemic Dangers of Context Collapse Online". In Lackey, Jennifer (ed.), *Applied Epistemology.* Oxford: Oxford University Press, 437-456.

Gillespie, Tarleton. (2018). Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions that Shape Social Media. New Haven: Yale University Press.

Introne, Joshua, Yildirim, Irem Gokce, Iandoli, Luca, DeCook, Julia, and Elzeini, Shaima. (2018). 'How People Weave Online Information Into Pseudoknowledge'. Social Media+Society, doi.org/10.1177/2056305118785639

Levy, Neil. (2017). 'Nudges in a Post-Truth World'. *Journal of Medical Ethics*, 43(8), 495-500.

Marwick, Alice E. (2018). 'Why Do People Share Fake News? A Sociotechnical Model of Media Effects'. *Georgetown Law Technology Review*, 474, 474-512.

Marwick, Alice E., and boyd, danah. (2011). 'I Tweet Honestly, I Tweet Passionately: Twitter Users, Context Collapse, and the Imagined Audience'. *New Media & Society*, 13(1), 114–133.

Miller, Boaz and Record, Isaac. (2013). 'Justified Belief in a Digital Age: On the Epistemic Implications of Secret Internet Technologies'. *Episteme* 10(2): 101-118.

—— (2017). 'Responsible Epistemic Technologies: A Social-Epistemological Analysis of Autocompleted Web Search'. *New Media and Society* 19(12) 1945-1963.

O'Connor, Cailin, and Weatherall, James O. (2019). *The Misinformation Age: How False Belief Spread*. New Haven: Yale University Press.

Origgi, Gloria. (2013). 'Democracy and Trust in the Age of the Social Web'. *Teoria Politica* 3, 23-38.

——(2018). *Reputation: What It Is and Why It Matters*. Princeton: Princeton University Press.

Phillips, Whitney, and Milner, Ryan M. (2021). *You are here: A field guide for navigating polarized speech, conspiracy theories, and our polluted media landscape*. MIT Press.

Priest, Susanna. (2013). Critical science literacy: What citizens and journalists need to know to make sense of science. Bulletin of Science, Technology & Society, 33(5-6), 138-145.

Record, Isaac. 2013. 'Technology and Epistemic Possibility'. *Journal for the General Philosophy of Science* 44: 319-336. Oxford: Oxford University Press.

Record, Isaac and Miller, Boaz. (2022). People, posts, and platforms: reducing the spread of online toxicity by contextualizing content and setting norms. Asian Journal of Philosophy. doi.org/10.1007/s44204-022-00042-2

Schwarz, Ori. (2019). 'Facebook Rules: Structures of Governance in Digital Capitalism and the Control of Generalized Social Capital'. *Theory, Culture & Society*, 36(4), 117-141.

Tufekci, Zeynep. (2017). *Twitter and Tear Gas: The Power and Fragility of Networked Protest*. Yale University Press.

Vaccaro, Kristen, Sandvig, Christian, and Karahalios, Karrie. (2020). '"At the End of the Day Facebook Does What It Wants": How Users Experience Contesting Algorithmic Content Moderation'. *Proc. ACM Hum.-Comput.Interact.4, CSCW2,* Article 167 (October), doi.org/10.1145/3415238

***Boaz Miller** is a visiting fellow at the Center for Advanced Internet Studies, Bochum, Germany and a senior lecturer at the Zefat Academic College Department of Management information Systems. He has a PhD from the IHPST, the University of Toronto. He works in social epistemology and philosophy of science and technology. His recent publications include "When is Scientific Dissent Epistemically Inappropriate?" Philosophy of Science (2021), and "Is Technology Value-Neutral?" Science, Technology, & Human Values (2021).*

***Isaac Record** is a visiting fellow at the Center for Advanced Internet Studies, Bochum, Germany and a teaching professor at Lyman Briggs College, Michigan State University. He has a PhD from the IHPST at the University of Toronto. He is the founding director of the Collaborative Experiential Learning Laboratory (CELL) and he teaches courses in philosophy of science, science and technology studies, and critical making. He researches how technology affects what we know. Find him at www.isaacrecord.net*