On Disinformation

The Internet Governance Project, a research center at the Georgia Institute of Technology's School of Public Policy, is pleased to submit its comments to the UN Special Rapporteur on freedom of expression. As advocates of free expression on the internet, we appreciate that the UN SR's call for comment recognizes that there is a tension between freedom of expression and efforts to crack down on disinformation.

Adopt a Narrow Definition

What is disinformation? Asking this question quickly reveals how problematic the whole concept is. The UN Special Rapporteur's call for comment defines disinfo as "false information that is created and spread, deliberately or otherwise, to harm people, institutions and interests." This definition contains a significant error. It fails to distinguish between false information that is spread *intentionally* to mislead or disrupt, and statements that may prove to be false but reflect uncertainty, differences of opinion, different belief systems, or disagreements about facts.

Any accurate and workable definition of disinformation must be confined to *intentional* dissemination of statements known to be false. Additionally, the term should be restricted to intentional falsehoods disseminated by actors with *the capacity to disrupt or strategically undermine authoritative information flows* in a community.

Contrary to the call for comments, which asserts that there is no uniform definition of disinformation, the definitions in the academic and professional literature shows a high level of congruence. They all classify disinformation as an *intentional* act. The High Level Expert Group on Fake News and Online Disinformation of the European Commission (2018, p. 3) defined disinformation as "all forms of false, inaccurate, or misleading information *designed, presented and promoted to intentionally cause public harm or for profit"* (emphasis ours). A study by Facebook defined disinformation as "Inaccurate or manipulated information content *that is spread intentionally*. This can include false news, or it can involve more subtle methods such as false flag operations, feeding inaccurate quotes or stories to innocent intermediaries, or knowingly amplifying biased or misleading information." (Weedon et al., 2017) Importantly, Facebook's definition adds that "Disinformation is distinct from misinformation, which is the inadvertent or unintentional spread of inaccurate information without malicious intent." Two communications scholars (Bennett & Livingston 2018) defined disinformation as "intentional falsehoods spread as news stories or simulated documentary formats to advance political goals." They also refer to it as "... systematic disruptions of authoritative information flows due to strategic deceptions ..."

Disinformation is thus strategic, intentional and, in the hands of powerful, well-resourced actors, potentially harmful. But it is not a common occurrence. Actors who are both powerful and capable of successfully utilizing strategic lies to create major social harms are relatively rare.

And free, competitive, diverse and open media systems are more resilient to manipulation than centralized and controlled ones.

Disinformation and political divides

The sudden targeting of “disinformation” has its roots in a backlash against social media that routinely lumps together “disinformation” with anything and everything that is considered objectionable about online communication. The term is now used interchangeably with:

- *computational propaganda* (which may or may not disseminate disinformation and could be nothing more threatening than tech-driven public relations campaigns),
- *misinformation* (which is just another word for erroneous assertions or beliefs),
- *conspiracy theories* (which might include anything from QAnon craziness to the Russian collusion charge against Trump)
- *hate speech* (which may just mean any expression that offends someone),
- *incitement* (which has a precise and narrow legal definition which may or may not rely on disinformation)
- *extremism* (which refers to a belief system not a form of disinformation.

Not all of this is disinformation by any definition, but all of it feeds into a growing narrative that online free speech is a threat to order and democracy, and mass surveillance and coercive policy measures must be taken to rein it in. This backlash is misguided.

The main problem with the existing discourse on disinformation is that it tends to conflate disinformation with any kind of false belief or questionable views. Equating the two invites systematic censorship. Because disinformation is considered actionable, classifying a contested or misinformed view as disinformation puts all dissent and disagreement at risk. Equating misinformation and disinformation is inherently inimical to freedom of inquiry, expression and thought; it is especially dangerous in a politically contentious environment. If the disinformation label includes any statements deemed to be false or misleading, competing political parties or movements can exploit power asymmetries to suppress or silence their opponents. They merely need to label them purveyors of disinformation. This is a much greater threat to democracy than disinformation.

Debates and disagreement over what is true, what is distorted, how to represent people or interpret events accurately are endemic to a liberal democracy. Panicking about disinformation betrays a fundamental lack of faith in the ability of citizens to govern themselves. Reaching for coercive tools to combat it are likely to worsen distortions and manipulations of public discourse, not make it better.

Most of the social problems attributed to “disinformation” are really problems of political and social division. Informational battles exploit and reflect those divisions; while they can make them worse, they do not create them. We agree with the insightful comment of Benkler, Faris and Roberts (2018) regarding the role of disinformation campaigns in American politics. They observed that U.S. susceptibility to disinformation campaigns:

> "does not come from Russia, though Russia clearly has been trying to exploit it. That susceptibility does not come from Facebook, though Facebook has clearly been a primary vector online. It comes from three decades of divergent media practices and consumption habits that have left a large number of Americans, overwhelmingly on the right of the political spectrum, ... ready to believe the worst, as long as it lines up with their partisan identity."

In fact, polarization makes the left side of the spectrum equally susceptible to informational manipulation. Responses to any disinformation problem, therefore, must be extremely careful not to make regulation in the name of combating disinformation a partisan political tool. A lodestar for any policy response is that social conflicts must be resolved through peaceful, democratic means, and freedom of expression is an essential component - indeed, the most important element - of the public's capacity to peacefully deliberate over what is true and what is a lie. To impose greater controls on speech and association simply gives the forces who wish to manipulate public discourse more powerful tools with which to do so.

Threat inflation

The UN call for comment exaggerates the risks of disinformation. It says "disinformation has a corrosive effect on democracy, development and human rights" and that it has "eroded public trust in democratic elections, contributed to incitement of violence and hatred, challenged sustainable development and undercut effective responses to the COVID-19 pandemic, endangering the lives of millions of people." No scientific evidence is provided to support these sweeping conclusions.

The most commonly cited basis for the claim of major societal harms is that Russian disinformation operations elected Donald Trump in 2016. But most serious researchers have rejected this claim (Benkler, Faris and Roberts, 2018; Rid, 2019).

The COVID-19 pandemic is also often cited as an example of the damage attributed to disinformation. In fact, it is an example of how the biggest problems we face in public communication are *not* attributable to disinformation. While it is true that a lot of confusing and false information has circulated about the virus, it is a major distortion of reality to portray public debates about pandemic policy as a simple case of true, authoritative information vs. intentional, strategic deception. Misunderstandings about covid, vaccines, masking and the like have resulted from high levels of scientific uncertainty, confusion, poorly organized messaging by public authorities, and - most importantly - legitimate disagreements about the social tradeoffs associated with lockdowns and re-openings. Wired Magazine reported in July 2020 that "In the early days of the pandemic, the World Health Organization, the Centers for Disease Control and Prevention, and even WIRED warned people against using masks." This included Dr. Fauci, who later admitted that he had engaged in a form of strategic deception to discourage mask use in order to prevent the public from depleting supplies for hospital workers. Six months later, "face coverings went from being discouraged by the world's top public health officials to being

encouraged by them—and from being opposed by US political leaders affiliated with the president to being accepted, if not demanded." Blaming these missteps in the global social response to the pandemic on "disinformation" or an "infodemic" rather than on uncertainty, trial and error, and divergent values and policy positions is simply incorrect. Note also that the government of China used a disinformation rationale to silence Dr. Li Wenliang, who issued one of the earliest alarms about the virus. He was suppressed for "disturbing social order" by "making false comments." Anti-disinformation actions can short-circuit the airing of crucial warnings and stifle legitimate differences over public policy and the interpretation of scientific evidence.

President Trump's claim that the 2020 election was stolen is another example of the alleged harms of disinformation. Yet it would be wrong to conclude that *disinformation* per se created the post-election crisis in America. All authoritative information sources and institutions - state and federal courts, state legislatures, state election officials from both parties, and nearly all mainstream media outlets - had rendered a negative verdict on Trump's assertion that he had won the election. True, Trump and his backers continued to assert, falsely, that the election had been stolen. But this was more an act of *propagandistic mobilization* than strategic deception. Trump's refusal to accept the results exploited deep cultural and political divisions between his supporters and his opponents in an attempt to hang on to power. No amount of speech regulation or content moderation could have prevented this. A sitting President and dozens of elected officials who supported him would have been able to air their message in one way or another. To see the Trump incident as a problem that was caused by disinformation ignores the political and social roots of the incident.

Ultimately, the problem caused by Trump's "big lie" was resolved politically, as it should have been. Democratic institutions and processes held. Trump and his Party were punished by voters. Trump not only lost the election, the Republican Party lost its Senate majority - due in no small part to Trump's attacks on the election's legitimacy in the state of Georgia, which energized Democratic voters and discouraged voters in his own party from participating in the runoff elections. The Capitol riots discredited the President further. Democracy and free expression proved to be more resilient than we are giving it credit for.

Appropriate countermeasures

The UN Special Rapporteur's call for comment asks pertinent questions about means of combating disinformation and their impact on human rights. In response, we begin by asserting key principles.

No policy or law can eliminate - at the source, before publication - all or even most forms of strategic deception to manipulate public discourse. Any law or policy based on the premise that it can, will by definition be authoritarian, because it will give a central authority rights of prior restraint over speech and make it the ultimate arbiter of truth. The risks of empowering a central authority with such wide-ranging powers over expression and inquiry far outweigh the risks of

disinformation. Free and democratic societies must rely primarily on public challenges to and debates. Based on those principles, responses to disinformation should:

- Be *ex post* or reactive, not based on prior restraint
- Rely primarily on public debate, exposure and contestation of false claims, and on publicizing counter-arguments, not suppression
- Leverage gatekeepers or authenticators who have earned public trust
- Be willing to allow its private media and platforms to suspend or block repeat, coordinated liars (after the fact)
- Rely on legal remedies against defamation or fraud as a last resort

Defamation lawsuits

Defamation is a liability regime that specifically targets factually wrong statements that are made knowingly and with harmful effect. Defamation law requires: 1) a false statement purporting to be fact; 2) publication or communication of that statement; 3) malicious intent or negligence; and 4) harm caused to the person or entity who is the subject of the statement. Note that defamation suits should only be directed at the perpetrators of the disinformation and not at the platform or intermediary who happens to transmit them.

The use of these laws by voting machine vendors who were libeled by the Trump movement has had immediate and beneficial effects on false claims about the U.S. election. Media outlets that were repeating baseless attacks on the vendors' voting technology were forced to issue disclaimers and retractions, and may yet be subject to financial penalties. When Alex Jones' Infowars ran stories asserting that yogurt maker Chobani was "Caught Importing Migrant Rapists" the company used defamation law and Jones was forced to settle. These kinds of legal remedies allow the truth or falsity of the charges to be formally adjudicated. Convictions and settlements act as an effective deterrent.

Section 230 and source recognition

Under the so-called "good samaritan" subsection of the law establishing platform immunity in the United States, platforms retain a wide latitude to suppress certain posts, suspend accounts, block retransmissions, or append disclaimers to posts. As long as these measures are taken after the fact (no pre-emptive or *ex ante* algorithmic filtering), and as long as there is diversity in social media platforms, and the measures are applied to a narrow band of the most disruptive or egregious cases, content moderation by platforms is a useful and appropriate way to curb disinformation. Content moderation can, however, be taken too far and moderation that is or even appears to be politically biased or partisan will do more to inflame social divisions than moderate them.

Account suspension or deplatforming as a remedy should be reserved for major threats and more well-resourced actors. Account suspension is justified for serial disinformers, including nation-state actors engaged in extended, repeated or coordinated disinformation campaigns.

Measures that produce human rights violations

Fake News Laws

It is now abundantly clear that legislative efforts to regulate disinformation - the so-called “Fake News Laws” passed or proposed in Malaysia, Brazil, Egypt, Indonesia and elsewhere - are for the most part nothing but censorship laws. They are used by governments to protect themselves from criticism or to harass or silence journalists. The Poynter Institute has a good but somewhat dated compilation of laws, task forces and policies around the world. Egypt is one of the worst examples. A law passed in July 2018 empowered the Supreme Media Regulatory Council to block websites and social media accounts with more than 5,000 followers for ‘fake news’ and can levy fines up to up to 250,000 Egyptian pounds ($14,400) without having to get a court order.

State-influenced Private Action

The growing intervention into social media content regulation by states and organized advocacy groups has created an area of “shadow regulation” in which private actors make the decisions but governments are looking over their shoulders. This combination destroys transparency and tends to promote politically biased content moderation. In India, for example, organizations and speech that might have been moderated or blocked based on Facebook’s policies were spared that treatment because they were affiliated with the ruling party. Facebook’s government affairs officer for India saw to it that the state would not be offended by its local content moderation decisions.

Conclusion

The most effective counterweight to disinformation is now and always has been to protect the freedom to challenge it, investigate its source and assert alternate points of view. While it is true that social media and other forms of technologically-driven communication have scaled up the connections among bad actors and the reach of some bad ideas, they have also by their very nature scaled up the connections and reach of good actors, useful information and benign movements. Exposures of police brutality, rapid dissemination of information about corporate or governmental malfeasance, advocacy and organization for social cooperation, new forms of fundraising and more efficient markets, and a growing awareness of public affairs are also well-documented consequences of the internet and social media. To see online communication as inherently dangerous and irrational is to reject liberal democracy itself. For if people are incapable of coping with sorting through the validity of the materials they find online, they are also not capable of governing themselves.