

Yuqing Zhai

09/18/2024

Vardi wrote an editorial in the communication for the ACM magazine criticizing the effects of information technology called *Computing: You Have Blood on your Hands*. He talked about the perpetual societal issues like hate speech youth mental health crises caused by social media platform. This essays will provide a brief summary Vardi's editorial and provide related policy vacuum and conceptual muddles in the issues. It will then analyze the ethical theories Vardi used and the potential logical fallacies, and argues that there is a lack of strong and direct evidence in part of his argument. **Using utilitarian ethics, this essay will argue that the online content censoring of misinformation and hate speech could have worser effect than not to.**

Vardi claims that creation of today's online social media platform, like Facebook and Twitter, have been used to *"hurt, traumatize, and even kill vulnerable people"* (ref) when in the absence of proper content moderation. He gives few instance illustrate this issue: (1) Hate speech prevalent on the Twitter after Musk fired Twitter's staff responsible for content moderation. (2) Anti-Rohingya content get amplified by Facebook's algorithm. (3) Facebook internal document shows its inability to control misinformation on the platform, potentially leading to the youth mental health crisis.

**He relies on both deontological and utilitarian ethics to show his points.** To give context, utilitarian ethics judges the morality of the action by the consequences it has caused, while deontological ethics judges actions based on one's intent to fulfill duty to others, emphasizing 'good will' as defined by Immanuel Kant. In Vardi's editorial, he notes the genocide done by Myanmar's security forces and Facebook's contribution in it, showing the **consequences of lack of content moderation**. At the end of the paragraph, he uses parallelism to show that tech industry's obsession with efficiency comes at the cost of resilience, alluding the **duty for tech professionals to provide secure software** and the immorality of not doing so as illustrated in the article.

**Vardi's argument suffers from few logical fallacies and inadequate evidences for some part of the reasoning.** In the beginning paragraph, he says *"Hate speech is prevalent on the Internet, and hate speech has consequences"*. (ref) **The word "hate speech" have been a conceptual muddles**, on the one hand, it could mean genuinely harmful, racist, or sexist comments, but on the other hands, it could be used to labeling other people's unpopular and controversial, who have expressed in good faith and with good reasons. Here, without proper clarification and more detailed evidence about the argument, he could fall into **fallacy of equivocation**, treating the *"hate speech"* with different meaning in his argument, one might be the quote one aforementioned.

Also, in his third example about Facebook's inability to moderate control misinformation, he mentions the "coincidence" between the tripled suicide rate among children from 2007 to 2017 and the release of iPhone and Facebook in 2007 and 2008. Although he refers Surgeon General who said "[technology could] undermine the safe and supportive environments young people need and deserve.", (ref) it is more of a coincidence in timing and a theory, and there is no direct scientific evidence to backup the causation between the misinformation in the platform and the mental health issue among teenager, and the degree of such causation if so. Without providing more detailed analysis here, **he might have treated correlation here as causation.**

**Content moderation, or content censorship might have a severe consequence.** As stated by the United State Constitution and the Universal Declaration of Human Rights, freedom of speech is essential right to every individual. Part of the reason is that without freedom of speech, government could do censorship to cover up tyrannical practices, and influence people's opinion in the favor of certain groups' interest, and there is the reason that Section 230 of Communication Act exists. It exempts the social media companies from the consequence of the content published by the users and allows platforms to remove content in "good faith". Therefore at least in United States, there is **currently no obvious policy vacuums** regarding to the content moderation done by the social media company, and section 230 are done to encourage the freedom of speech on digital platform.

From the twitter files released by Elon Musk on the twitter, we see that there might be involvement of FBI in forcing twitter to censor certain content including election, which is a direct violation of constitution if the case is settled. (ref) . There are pressures from BioNTech along with German government for twitter to censor opinion against of general COVID vaccines, mainly for the interest of profit. (ref) Both cases could deeply manipulate the people's opinion on social affairs and affect their decision making that might have greater consequence on the country (election example) and their family (vaccine example).

The point here is not against the moderation of certain content in 'good faith' genuinely reduce the harm it might bring to individuals, as Vardi says, but to be aware of the possibility, that **if the censorship is done in the name of "removing misinformation and hate speech", the consequence from a tyrannical government might be way more severe than the hate speech from other individuals.**

**Reference**

<https://cacm.acm.org/opinion/computing-you-have-blood-on-your-hands/>

[https://en.wikipedia.org/wiki/Twitter\\_Files](https://en.wikipedia.org/wiki/Twitter_Files)

<https://theintercept.com/2023/01/16/twitter-covid-vaccine-pharma/>

<https://www.un.org/en/about-us/universal-declaration-of-human-rights>