2-26-2020

The clash of empires: Regulating technological threats to civil society

Tracey Leigh Dowdeswell

Nachshon Goltz
Edith Cowan University

10.1080/13600834.2020.1735060
This is an Accepted Manuscript of an article published by Taylor & Francis in *Information & Communications Technology Law* on 26 February 2020, available online: http://www.tandfonline.com/10.1080/13600834.2020.1735060
Dowdeswell, T. L., & Goltz, N. (2020). The clash of empires: Regulating technological threats to civil society. *Information & Communications Technology Law, 29*(2), 194-217. https://doi.org/10.1080/13600834.2020.1735060

The Clash of Empires: Regulating Technological Threats to Civil Society

By Tracey Leigh Dowdeswell* and Nachshon (Sean) Goltz**
*Faculty, Humanities & Social Sciences, Douglas College (Canada)
**Senior Lecturer, School of Business & Law, Edith Cowan University (Australia)

Abstract

This paper examines the regulation of technology platform companies – those companies that provide a platform for user-generated media content. These companies play an increasingly dominant role in the global flow of news and information. In doing so, platform companies play a crucial role in modern civic life, by deciding which content will reach users, engage the public's attention, and be deemed credible. It is therefore crucial that we choose means of regulation that foster democratic values and robust civic engagement. In this paper we focus on the regulation of 'computational propaganda,' including misinformation and 'fake news,' the rise of synthetic media and so-called 'deep fakes,' and novel forms of algorithmic injustice, such as the manipulation of search engine results and their effect on elections. We argue that many existing regulations fall short in that they adopt an approach that views regulation as a battle between two competing powers, or 'empires' – that of the regulatory state versus the big tech companies. Accordingly, they approach regulation as a means of redistributing power between these two players, while discounting the end user, and they often involve unjustified restrictions of free speech through the imposition of content controls. We propose instead three guidelines for platform governance that serve as goals or guideposts to regulating computational propaganda without the need for content controls, these being: 1) transparency – about the sources of information, their funding, and their credibility, 2) the promotion of decentralization and user control over flows of information, and 3) efforts to enhance media literacy and credibility. We need to choose methods of regulating technology that foster democratic rights without empowering the kind of censorship that goes along with censorship and content controls.

Keywords: Regulation of technology, information governance, media literacy, computational propaganda, fake news, synthetic media, technology and civil society.

1. Introduction: Regulation & The Clash of Empires

The Mongol Empire laid siege to Baghdad in 1258 A.D., overthrowing the Abbasid Caliphate, and sacking the city that had been the centre of Islamic cultural life for very nearly five hundred years. To impose the rule of the new empire, the Mongols first destroyed the trappings of the old: they executed its Caliph and eliminated its royal family; they decimated its armies and its population; they tore apart the books in its libraries and destroyed the august House of Wisdom. There were stories of Mongol soldiers using the vellum torn from priceless manuscripts for their sandals.[1] The Mongols sought to dominate the inhabitants by outlawing kosher and halal butcheries, and they prohibited the people from refusing to eat Mongol foods.[2]

But the wholesale destruction of the old empire by the new was not – as it never is – the whole story. The new regime needs to incorporate portions of the old to consolidate its power. The technologically advanced Mongol Empire quickly came to dominate Central Asia: the introduction of stirrups, the improvement of light cavalry, the cultivation of horse milk as a source of nutrition all gave Mongol horsemen the advantage on long campaigns across the steppes. But the new empire had little experience in collecting taxes, maintaining order, or administering a cosmopolitan empire with its great bureaucracy and far-flung possessions.

The Mongol Khan of Soltaniyeh – located near the Caspian Sea in northern Iran – needed to employ Persian scholars to help him administer his corner of the empire. The chief of these was Rashid Al-Din, a highly educated scion of the old elite. Al-Din was a cosmopolitan scholar and politician of international renown. Originally from a Persian Jewish family, he converted to Islam and rose to become the Grand Visier of the Mongol Khanate. The story is told that the

[1] Stewart A. P. Murray, *The Library: An Illustrated History* (New York: Skyhorse Publishing), 54.
[2] Johan Elverskog, *Buddhism and Islam on the Silk Road* (Philadelphia: University of Pennsylvania Press, 2010), 228.

second Khan, Mahmoud Ghazan, not only made Al-Din kneel before him throughout their meetings, but would make him stand behind his chair during state banquets; if the Visier's service pleased him, Khan Ghazan might toss him a piece of meat (non-Halal, of course) over his shoulder.[3]

What does this historic *transregnum* tell us about our present situation? As in the past, we are faced with a rising power that is technologically sophisticated, but lacks an administrative architecture. The new power therefore depends a great deal on the expertise, the cooperation and even the complicity of the existing elite at the same time as it seeks their domination – and occasional humiliation. There are lessons in this for our own time, as we seek to use old political and legal structures to regulate the rising power of technology platform companies and their influence over civic life. We argue that it is wise to see these as two facets of the same ecosystem rather than competing powers. Our task is therefore not to distribute power between them, but to manage the broader econsystem of which they are a part, and so to promote its most optimal functioning.

In our own time, this ecosystem is not a physical empire – with its possessions, its tax collectors and its armies – but the flows of information in a technological and highly complex digital landscape. The Mongols sought to control their new subjects through the destruction of the old culture and its values, and imposed controls over what kinds of foods the subjects could not eat – even going so far as to regulate what foods they *must* eat. This is a metaphor for the imposition of content controls on media in our own time – a time in which the shift in power and the governance task at hand is equally monumental. What is often lost in the broader debates over what information people should and should not access is the essential dignty of the end user

[3] John Glubb, *Fate of Empires and Search for Survival* (Edinburgh: William Blackwood & Sons, 1977), 11-12.

and their freedom to choose. Here, we place ths principle at the heart of our regulatory regime, and instead seek to govern platform companies in the manner that will best promote civic values and user choice.

We begin with a brief examination of the U.S. anti-trust case against Microsoft as an example of the clash of the old and the new empires, one in which the struggle for power between the great players takes precedence over the struggle for democratic values and individual autonomy. We then look at three new technological domains that have the power to either strengthen or subvert the regulatory state, depending on the means we choose to regulate them. Together these are examples of what might be termed ‘computational propaganda,’ including misinformation and ‘fake news,’ the rise of synthetic media and so-called ‘deep fakes,’ and novel forms of algorithmic injustice, focusing specifically on the manipulation of search engine results and their effect on election results. We then examine various methods of governance to deal with these new technological challenges, to better determine which will empower the big players, and which will empower human beings – as users of technology, as consumers of media, as citizens of democratic states, as holders of rights, and as practitioners of civic values.

In the final section, we propose three governance domains that serve as goals or guideposts to regulating computational propaganda without the need for content controls, these being: 1) transparency – about the sources of information, their funding, and their credibility, 2) user control over information and its decentralization away from large organizations, and 3) media literacy and efforts to enhance our ability to distinguish the credibility of the information we are consuming. We need to choose methods of regulating technology that foster democratic rights without empowering the kind of censorship that goes along with censorship and content controls, lest we end up as spectators in a twenty-first century clash of empires, fighting for the scraps left under the table.

2. Technology Platform Companies & Civic Life

Technology companies provide platforms for many of the services we now need to participate in a modern economy. They control the global flow of information, they provide news, content, communication, culture, and social interaction for a large and increasing portion of the world's peoples. The largest of the platform companies – Apple, Amazon, Google (Alphabet) and YouTube, Facebook and its subsidiaries Instagram and WhatsApp, as well as Twitter – hold a near-monopoly share of their markets.[4] About 90% of all searches globally are done through Google; YouTube and Facebook video hold a majority – at about 89% – of the desktop online video market, while Facebook holds a majority of all social media log-ins.[5]

Technology companies wield great economic power, and they dominate the shifting boundaries of our security and privacy in the digital sphere.[6] As Moore states:

> These global technology giants now bestride our world like Colossi. We wake to their alarms. We sleep to the continual ping of messages arriving on their hardware and via their software. They have become integral to our communication, to our access to news and information, to our virtual identities.[7]

Platform companies monitor and collect information about our online activities – Facebook tracks its users' online activities even at times when they are not using its platform[8] – all for the purpose of controlling and directing our behaviour to line up with corporate interests.

[4] Martin Moore, "Tech Giants and Civic Power," Centre for the Study of Media, Communications and Power, King's College London (April 2016), 3.
[5] Moore, "Tech Giants and Civic Power," 13.
[6] Moore, "Tech Giants and Civic Power," 3.
[7] Moore, "Tech Giants and Civic Power," 10.
[8] Moore, "Tech Giants and Civic Power," 10.

At the same time, the digital empire is closely aligned with the old empire of the nation state, and its bureaucratic and administrative apparatus. It supports state institutions at the same time as it relies on them for its legal and economic privileges. Some argue that the current U.S. President was elected with the help of media misinformation delivered on platform companies like Facebook;[9] others have voiced concerns about the power of America's leading technology czar – and the richest man in the world – when Jeff Bezos of Amazon closed a deal to provide $600 million worth of cloud computing to the U.S. Central Intelligence Agency.[10] Meanwhile, Google has been accused of supporting Chinese military interests over those of America.[11] Platform companies have been accused of fostering the surveillance capabilities of illiberal regimes around the world and oppressing human rights.[12]

Platform companies are now performing civil functions we would normally associate with state power. Facebook and Twitter have taken over some communications functions that are *explicitly* civic, including official communications during natural disasters, amber alerts, and extreme weather warnings.[13] More than this, they now possess the "power to enable collective action, the power to communicate news, and the power to influence people's vote."[14] But as Woolley and Howard point out, platform companies that once held out the promise of

[9] Hannah Jane Parkinson, "Click and Elect: How Fake News Helped Donald Trump Win a Real Election," *The Guardian*, November 14, 2016, https://www.theguardian.com/commentisfree/2016/nov/14/fake-news-donald-trump-election-alt-right-social-media-tech-companies.

[10] Norman Solomon, "If Obama Orders the CIA to kill a U.S. Citizen, Amazon will Be a Partner in Assassination," *Alternet*, February 12, 2014, http://www.alternet.org/print/news-amp-politics/if-obama-orders-cia-kill-us-citizen-amazon-will-be-partner-assassination.

[11] Zak Doffman, "Google Accused by Top U.S. General and Senator of Supporting Chinese Instead of U.S. Military," *Forbes*, March 16, 2019, https://www.forbes.com/sites/zakdoffman/2019/03/16/google-accused-by-u-s-general-and-senator-of-benefiting-chinese-instead-of-u-s-military/#1b5b83df1899

[12] See eg., Joe Odell, "Inside the Dark Web of the UAE's Surveillance State," Middle East Eye, February 27, 2018, https://www.middleeasteye.net/opinion/inside-dark-web-uaes-surveillance-state. This describes the company "Dark Matter", and its close alliance with the UAE government to engage in mass surveillance, including accessing and blocking users' social media content.

[13] Moore, "Tech Giants and Civic Power," 27.

[14] Moore, "Tech Giants and Civic Power," 4.

empowering users and fostering political action have betrayed this by their increasingly illiberal tendencies, stating that “social media can be used very effectively for propaganda and political control; surveillance capacities are running way ahead of our ability to catch up with civil liberties protections; and filter chambers are becoming entrenched and are sowing extremist beliefs and social divisions.”[15]

Google’s algorithms decide what web pages take precedence at the top of the search rankings, and which ones will be buried on page 2 – and therefore nearly never clicked on. Amazon decides what books receive precedence in searches, which intellectual works are to be promoted and recommended to purchasers, and which will not. Facebook decides which news stories will be showcased in their Newsfeed, and which will be buried as ‘fake news.’ In the same way that the printed press and the nascent field of journalism altered the nature of governance and democracy from the eighteenth to twentieth centuries, so too will the digital communications landscape alter it for the twenty-first.[16] This time, the players are much more powerful, more centralized, have a wider reach, and there are fewer of them.

Moore provides a list of six ways in which tech companies are exercising civic power: 1) the power to command the public’s attention; 2) the power to communicate news and information; 3) the power to enable collective action; 4) the power to give people a voice; 5) the power to influence people’s vote; 6) the power to hold other powers to account.[17] These companies play a dominant role as gatekeepers to information, and hence to public opinion – and even voting behavior. At the same time, novel forms of computational propaganda, combined

[15] Samuel C. Woolley and Philip N. Howard, “Introduction: Computational Propaganda Worldwide,” in *Computational Propaganda: Political Parties, Politicians, and Political Manipulation on Social Media*, ed. Samuel C. Woolley and Philip N. Howard (Oxford: Oxford University Press, 2019), 3.
[16] Moore, “Tech Giants and Civic Power,” 26.
[17] Moore, “Tech Giants and Civic Power,” 24.

with the power and reach of the platform companies that distribute them, pose particular risks to these six domains of civic life. This in turn poses novel and distinct threats to democracy, and this is one of the greatest governance challenges we face in the coming decades.

The methods we adopt to govern platform companies need to be chosen carefully if they are to promote democratic values and civic engagement. In controlling the flow of information, there is a risk that we will instead empower the state or its corporate proxies – through oppressive policies that violate free speech, enable government surveillance, and quash essential freedoms like debate and association. In this paper, we examine the regulation of technology platform companies through the lens of the clash of empires – between the technological empire of the platform companies, new media and Silicon Valley on the one hand, versus the old empire of the regulatory state and its administration on the other – to better identify regulatory responses that will empower citizens, and which will merely feed back into the power struggles between the tech giants and the state.

3. Microsoft: Monopoly & The Clash of Empires

In this section, we will briefly examine the U.S. anti-trust case against Microsoft as an example of the clash of the old and the new empires, in which the struggle for power between the great players takes precedence over the struggle for civic values and user choice. This will hold important lessons, as we examine the unique challenges that modern technologies are posing for democratic governance.

The dominance of the technology empire globally is described as, "an almost unimaginable power to determine what we see, where we spend, how we perceive."[18] Legal scholars have

[18] Frank Pasquale, *The Black Box Society: The Secret Algorithms that Control Money and Information* (Boston: Harvard University Press, 2016), page 98.

analogized this power to the rentier economy of medieval feudalism,[19] in which “a small cadre of the lucky, the talented, and the ruthless are reaping huge returns from content put out by others.”[20] From its capital in Silicon Valley, the technology empire controls the online world and its most valuable resource – information. It even has its own cryptocurrency and, some argue, its own religion.[21]

Comparing wealth and ‘population’ between the two empires shows the striking rrach of the new one. Amazon’s market value is now about $797 billion, which would make it the 146th richest country if it were a nation state.[22] Google has a market value of about $748 billion and controls 92% of the world’s search market.[23] Facebook counts “one quarter of humanity as monthly active users,”[24] and has a market value is $479 billion,[25] not counting Instagram and Whatsapp which it also owns.[26] In comparison, the U.S. has a debt of $70 trillion and population of only 500 million.[27]

Google’s Revenue for 2018 was $136B, with almost 100K employees, bringing its Gross National Income (GNI) per capita to $1.36M. Facebook has 40K employees and its revenue for

[19] Greg Lastowka, *Virtual Justice: The New Laws of Online Worlds* (New Haven, CT: Yale University Press), 153, discussing whether “virtual societies are premised on owner-user relationship similar to those between lords and feudal societies”); Bruce Schneier, “Feudal Security,” Schneier Blog, December 3, 2012, http://www.schneier.com/blog/archives/2012/12/feudal_sec.html.

[20] Jaron Lanier, Who Owns the Future? (New York: Simon & Schuster, 2014), page 99.

[21] Douglas Rushkoff, “The Anti-Human Religion of Silicon Valley,” Wired, December 12, 2018, https://medium.com/team-human/the-anti-human-religion-of-silicon-valley-ac37d5528683.

[22] Much Needed, “Amazon by the Numbers: Stats, User Base & Fun Facts,” Much Needed Blog (2018), https://muchneeded.com/amazon-statistics/.

[23] Lauren Feiner, “Amazon is the Most Valuable Public Company in the World After Passing Microsoft,” *CNBC News*, January 7, 2019, https://www.cnbc.com/2019/01/07/amazon-passes-microsoft-market-value-becomes-largest.html.

[24] Gabriel Kahn, “Facebook’s Self-Deception,” *International Policy Digest*, January 25, 2018, https://intpolicydigest.org/2018/01/25/facebook-s-self-deception/.

[25] MarketWatch, “Facebook Worth $45 Billion More After Earnings Set it Up to Be Potential ‘Comeback Story of 2019,’” *MarketWatch*, January 31, 2019, https://www.marketwatch.com/story/facebook-may-be-the-comeback-story-of-2019-as-stock-soars-higher-2019-01-31.

[26] Dan Noyes, “The Top 20 Valuable Facebook Statistics – Updated July 2019,” Zephoria Digital Marketing Strategic Insights, Blog, July 2019, https://zephoria.com/top-15-valuable-facebook-statistics/.

[27] https://edition.cnn.com/2019/07/17/investing/united-states-debt-risks/index.html

2018 was $55B bringing bringing its GNI per capita to $1.37M.[28] Amazon revenue was $233B with 647K employees bringing its GNI per capita to $360K.[29] At the same time, the American empire GNI per capita was only $63K. In fact, the tech empire's combined GNI per capita is larger than the combined GNI per capita of the *top 107 countries in the world*.[30]

Struggles for dominance between these two empires are not new. The U.S. 2001 anti-trust action against Microsoft was an early battle between the more defensive-minded of the old empire, and one of the most aggressive and profitable companies of the new.[31] In this case, the U.S. government accused Microsoft of illegally maintaining its monopoly position in the personal computing market primarily through the legal and technical restrictions it put on the abilities of manufacturers and users to uninstall Internet Explorer and use other programs such as Netscape and Java.

In November 5, 1999 Judge Thomas Penfield Jackson issued his initial findings of fact, finding that Microsoft held monopoly power and used it to harm consumers, rivals, and other companies. Judge Jackson further ordered the break-up of Microsoft into two companies. Nonetheless, in June 28, 2001 a federal appeals court reversed the breakup order.[32]

There can be no doubt as to who won this battle. The case had little effect on Microsoft's behaviour. The fines, restrictions, and monitoring that the government imposed on the company were not enough to prevent it from further "abusing its monopolistic power."[33] This is no

[28] Daniel Priestley, Learn the Simple Equation That Tells You If Your Business Will Grow and Scale, (Sep 18, 2019), Entrepreneur, https://www.entrepreneur.com/article/337745

[29] J. Clement, Annual Net Revenue of Amazon from 2004 to 2019, (Feb 3, 2020), Statista, https://www.statista.com/statistics/266282/annual-net-revenue-of-amazoncom/

[30] List of Countries by GNI (PPP) per capita, Expedia, https://en.wikipedia.org/wiki/List_of_countries_by_GNI_(PPP)_per_capita

[31] *United States* v. *Microsoft Corporation*, 253 F.3d 34 (D.C. Cir. 2001).

[32] U.S. v. Microsoft: Timeline, WIRED NEWS, Nov. 4, 2002, http://www.wired.com/techbiz/it/news/2002/11/35212

33. Gregory T. Jenkins and Robert W. Bing, "Microsoft's Monopoly: Anti-Competitive Behavior, Predatory Tactics,

surprise considering the architects of this legal arrangement and their approach towards the technology empire online territory.[34]

Judge Posner, the chief mediator in the Microsoft case, later rendered an important decision in the legal action taken against the online sex-trafficking website Backpage.com.[35] In the *Backpage* case, Posner ruled in favour of the company, stating "What about old men who like to be seen with a young woman, right. That is an aspect of escort service, it's not all sex."[36] Judge Posner conceived of the problem as a struggle between two great and competing powers: corporate freedom of action and expression versus the government's power to regulate and control that freedom. His task, then, was merely to pick which side would win, but without reflecting that approaching the problem in this way missed the opportunity to balance competing rights and responsibilities for technology users and citizens in a democratic polity. This is a clear example of legal actors empowering a company at the expense of the government and its power to regulate corporate actors in the interests of a vulnerable group.[37] In this case, too, the company won, but what was lostwere the interests of the women and girls who were sexually exploited by the website and its customers.

If we step out of the mindset of viewing technology regulation as a clash of empires – as a choice between greater freedom for corporations to earn profits or greater power for the regulatory state – we gain a focus on the needs of technology users, civic actors, the needs of the vulnerable, and the values that will best foster human rights and democratic practices.

and the Failure of Governmental Will," *Journal of Business & Economics Research* 5 (2007): 222.

[34] Chris Butts,The Microsoft Case 10 Years Later: Antitrust and New Leading "New Economy" Firms, 8 Nw. J. Tech. & Intell. Prop. 275 (2010). https://scholarlycommons.law.northwestern.edu/njtip/vol8/iss2/5

[35] https://www.iamjanedoefilm.com/

36. *Backpage.com LLC* v. *Tom Dart, Sherriff of Cook County, Illinois*, No. 15-3047 U.S. 7th Circuit Court of Appeals, November 2015.

[37] Professor Lessig serving as "special master" in the Microsoft case, defended Google when it was described as a parasite extracting value created by others, saying that this is much like, "Leonardo da Vinci was just a 'parasite' upon the hard work of the paint makers."

Empowering the state to regulate by suppressing freedoms and ignoring the vulnerable is a road that has long led to oppression and social injustice. On the other hand, empowering corporations can be equally oppressive, socially dysfunctional, and economically wasteful. As one former Facebook executive said, "the best minds of my generation are thinking about how to make people click ads."[38]

4. Computational Propaganda: New Threats to Civic Society

4.1 Misinformation:

In this section, we describe three types of computational propaganda and how they are posing novel and as-yet unregulated threats to democratic values and civil society. We focus on three categories of computational propaganda that have been shown to pose serious and immediate risks: misinformation, synthetic media, and the search engine manipulation effect (SEME).

Computational propaganda encompasses much more than so-called 'fake' or 'junk' news. Woolley defines computational propaganda as "the use of algorithms, automation, and human curation to purposefully manage and distribute misleading information over social media networks."[39] The European Commission has defined disinformation as "all forms of false, inaccurate, or misleading information designed, presented and promoted to intentionally cause public harm or for profit."[40] Fake news publications are but one form of computational propaganda, and have been defined as those news stories that are "knowingly and intentionally

[38] Gloria Liou, "This is Silicon Valley," *OneZero*, February 27, 2019, https://onezero.medium.com/this-is-silicon-valley-3c4583d6e7c2.

[39] Samuel C. Woolley and Philip N. Howard, "Introduction: Computational Propaganda Worldwide," in *Computational Propaganda: Political Parties, Politicians, and Political Manipulation on Social Media*, ed. Samuel C. Woolley and Philip N. Howard (Oxford: Oxford University Press, 2019), 4.

[40] European Commission, "A Multi-Dimensional Approach to Disinformation" Report of the Independent High Level Group On Fake News and Online Disinformation (March 2018), 5.

false;"[41] this excludes satire, as well as news stories that may not be accurate, but are sloppily produced rather than intentionally misleading.[42] Computational propaganda tends to reference notorious public figures or controversial topics with the intention of generating clicks and going 'viral.'[43] Fake news and disinformation are devised to capture attention – to generate views, clicks and advertising revenue – in an information ecosystem in which sensationalism, novelty, and revenues are the main incentives for communication, and not accuracy and civic responsibility.[44]

Computational propaganda includes not only the production of false or misleading information, but also its distribution over the digital landscape and the context in which it is consumed and interpreted. Computational means of producing propaganda are in many ways a new iteration of an old practice: propaganda has long included communications that "deliberately subvert symbols, appealing to our baser emotions and prejudices and bypassing rational thought, to achieve the specific goals of its promoters."[45] Digital communication is more powerful than older methods of propagandizing: it is cheaper to produce and more efficient to spread, while promoting much higher levels of personalization and engagement with users. At the same time, computational propaganda is hidden, automated, and largely anonymous, making these methods much more challenging to regulate."[46] As referred to this in the artificial media laws theory, the more engaging is the particular media, the more it will take over the imagination and cognitive reasoning of the user, rendering them less able to judge the credibility of the content.[47]

[41] 41. David O. Klein and Joshua R. Wueller, "Fake News: A Legal Perspective," *Journal of Internet Law* 20, no. 10 (2017): 6.

[42] 42. Klein, "Fake News: A Legal Perspective," 6.

[43] 43. Klein, "Fake News: A Legal Perspective," 6.

[44] Oreste Pollicino and Elettra Bietti, "Truth and Deception Across the Atlantic: A Roadmap of Disinformation in the U.S. and Europe," *Italian Journal of Public Law* 11, no. 1 (2019): 46.

[45] Woolley, "Introduction: Computational Propaganda Worldwide," 5.

[46] Woolley, "Introduction: Computational Propaganda Worldwide," 7.

[47] Nachshon Goltz, and Tracey Dowdswell, *The Imaginationless Generation: Lessons from Ancient Culture in*

Computational propaganda also employs wholly novel methods of producing and distributing misinformation. One such method is the use of automated systems, known as ‘bots.’ Bots are particularly engaging and persuasive – many can even have seemingly genuine conversations with users – and so they are very effective at promoting or discrediting political messages and policy positions.[48] Bots can work to create or destroy an illusion of popularity and consensus – manipulating to their advantage our inborn tendency to follow the herd.[49] In this way, they are very effective in both facilitating *and* impeding political organizing. The related practice known as ‘astroturfing’ uses this same principle to create fake grassroots movements that appear genuine and highly popular[50] – all for the purposes of political persuasion, changing votes, and “nudging” mass behaviour.

The spread and consumption of misinformation through platform companies is as crucial a problem as its actual content. As of 2017, 25% of all Americans obtained their news from social media sites, up substantially from 15% just four years earlier.[51] Facebook and Google in particular are now major news platforms; about 40% of Americans are getting their political news from Facebook and about a third directly from Google.[52] Despite this, we are not able to assess how these companies filter and rank news stories, as these are done by proprietary and opaque algorithms.[53] As Moore states, “these, and other emerging services, give these information intermediaries a crucial role in determining what news citizens are exposed to (or

Regulating New Media (Lieden: Brill, 2019).ß

[48] Woolley, “Introduction: Computational Propaganda Worldwide,” 4.

[49] Woolley, “Introduction: Computational Propaganda Worldwide,” 9.

[50] Woolley, “Introduction: Computational Propaganda Worldwide,” 10.

[51] Elizabeth Greico, “More Americans are Turning to Multiple Social Media Sites for News,” *Pew Research Center*, November 2, 2017, https://www.pewresearch.org/fact-tank/2017/11/02/more-americans-are-turning-to-multiple-social-media-sites-for-news/.

[52] Moore, “Tech Giants and Civic Power,” 31.

[53] Moore, “Tech Giants and Civic Power,” 33.

not exposed to), how diverse this news is, and how it is prioritized and filtered."[54]

On Facebook, the kinds of news stories people receive in their feed depends to a good extent on who their friends are; this can promote what are called 'filer bubbles' and 'echo-chambers', in which people are exposed to a small number of viewpoints that tend to confirm the opinions they already hold.[55] As Epstein states, platform companies foster selective exposure, which leads to "situations where people become trapped in a digital echo chamber of information that confirms and strengthens their existing beliefs."[56] Stories are recommended to readers based upon their demonstrated habits and interests, leading them to receive more of the same kinds of stories. Despite the massive proliferation of information, the consolidation of media power by the tech companies means that people are actually receiving *less* information from *fewer* news sources.[57] At the same time, there is less funding for original and investigative journalism, which leads to fewer and lower quality news stories.

Filter bubbles are a civic danger in and of themselves, no matter the actual content of the news they contain, or whether the information they amplify and transmit is false. Part of the reason for this is the highly emotive and polarizing nature of junk news. Marchal found that junk news stories on Facebook were less prolific but more engaging, earning from 1.2 to 4 times as many likers and shares as credible news stories; this increased engagement is likely due to the highly emotive and often anger-provoking nature of the stories.[58] Filter bubbles also reinforce

[54] Moore, "Tech Giants and Civic Power," 32.

[55] Moore, "Tech Giants and Civic Power," 33.

[56] Robert Epstein, Ronald E. Robertson, David Lazer, and Christo Wilson, "Suppressing the Search Engine Manipulation Effect (SEME)," *Proceedings of the ACM on Human-Computer Interaction* 1, no. 2 (2017): 4.

[57] Elizabeth Greico, "More Americans are Turning to Multiple Social Media Sites for News," Pew Research Center, November 2, 2017, https://www.pewresearch.org/fact-tank/2017/11/02/more-americans-are-turning-to-multiple-social-media-sites-for-news/.

[58] Nahema Marchal, Bence Kollanyi, Lisa-Maria Neudert, and Philip N. Howard, "Junk News During the EU Parliamentary Elections: Lessons from a Seven-Language Study of Twitter and Facebook," Oxford Internet Institute, Data Memo 2019.3, https://comprop.oii.ox.ac.uk/wp-content/uploads/sites/93/2019/05/EU-Data-Memo.pdf, 4.

partisanship, which social media has pushed to unprecedented levels. As Khan states, “this certainty in the righteousness of our own point of view makes us regard a neighbor with a yard sign the way a Capulet regards a Montague. It seems to me that we suddenly hate each other a whole lot more than we ever did before.”[59] Uncivil media makes civil hands unclean.

Platform companies can have an outsized influence over not only the content that people are reading, but even over their ability – or not – to organize for political causes. As Moore states, “these organisations can give people a voice they can also take it away and, as commercial organizations, they do not need to say why. Given the dominance of these services, going to another service is often not a viable option.”[60] Ethan Zuckerman points out that these are private platforms masquerading as public forums, which holds serious consequences for their increasing monopoly over political news and political communication. He states:

> Hosting your political movement on YouTube is a little like trying to hold a rally in a shopping mall. It looks like a public space, but it’s not – it’s a private space, and your use of it is governed by an agreement that works harder to protect YouTube’s fiscal viability than to protect your rights of free speech.[61]

People and channels that express unpopular views can easily be ‘de-platformed’ from the public debate by private companies, either by limiting revenue streams or out-and-out content removal.[62] This can be done with little transparency or recourse and, even if the service is later restored, the company’s action may have already altered the terms of the public debate at an opportune moment.

[59] Gabriel Kahn, “Facebook’s Self-Deception,” International Policy Digest, January 25, 2018, https://intpolicydigest.org/2018/01/25/facebook-s-self-deception/.

[60] Moore, “Tech Giants and Civic Power,” 44.

[61] Ethan Zuckerman, “Public Spaces, Private Infrastructure – Open Video Conference,” My Heart’s in Accra, October 1, 2010, http://www.ethanzuckerman.com/blog/2010/10/01/public-spaces-private-infrastructure-open-video-conference/.

[62] Glenn Harlan Reynolds, “When Digital Platforms Become Censors,” *Wall Street Journal*, August 18, 2018, https://www.wsj.com/articles/when-digital-platforms-become-censors-1534514122?mod=rsswn.

There is a growing sense, particularly during election campaigns, that junk news – information that is "ideologically extreme, misleading, and factually incorrect" – is widespread and destructive of democratic functioning.[63] As Neudert states, "[i]n Europe, groups at the fringe of the political spectrum successfully used fabricated falsehoods, conspiracy theories, and hate speech to mobilize voters and sow public discontent with political systems."[64] Because these platforms are now so large, it becomes all-too-easy for a platform to bury almost any needle into a trillion-byte haystack.[65] Because these forms of information management are less visible, they can 'nudge' citizens into behaviour and beliefs that are more conducive to the authorities' wishes but without provoking the anger or backlash that overt censorship would do.[66]

Misinformation can be a very powerful influencer of people's behaviour and opinions. The large platform companies have the ability to control what information people can access, interpret, and engage with, and this in turn leads to changes in opinions, behaviours, and even votes.[67] Computational propaganda has been identified by the World Economic Forum as one of the top ten threats to democratic societies.[68] Because many of these methods of spreading computational propaganda are new, we have yet to fashion the legal or governance mechanisms to come to terms with them. Despite the power that platform companies now command over the news, they are not accountable to the same laws and regulations that constitutions and national laws wield over regular media.[69] Yet this is an area in which the dangers of state regulation are

[63] Marchal, "Junk News During the EU Parliamentary Elections," 1.

[64] Lisa-Marie N. Neudert, "Germany: A Cautionary Tale," in *Computational Propaganda: Political Parties, Politicians, and Political Manipulation on Social Media*, ed. Samuel C. Woolley and Philip N. Howard (Oxford: Oxford University Press, 2019), 153.

[65] Moore, "Tech Giants and Civic Power," 49.

[66] Margaret E. Roberts, *Censored: Distraction and Diversion Inside China's Great Firewall* (Princeton: Princeton University Press, 2018), 8-9.

[67] Woolley, "Introduction: Computational Propaganda Worldwide," 5.

[68] Woolley, "Introduction: Computational Propaganda Worldwide," 13.

[69] Kahn, "Facebook's Self-Deception."

equally great, as increasing state control over private communications may foster authoritarianism, and suppress free speech and open debate.

4.2 Synthetic Media:

Synthetic media – sometimes referred to as 'deep fakes' – use the power of machine learning and neural networks to create synthetic audio, photo, video, or even text media that appear authentic. One common example is video footage that has been altered to transpose one person's speech, mannerisms, and movements onto that of another.[70] Synthetic media now also include false e-mails, texts or messages, some of which appear to be from friends who know you well.[71] Currently, video technology can only make changes to the face beneath the forehead; it cannot yet replace an entire body, but programmers are working on this, and the technology is progressing.[72] Synthetic media is developing faster than technologies that detect it, making it nearly invisible and highly manipulative.

Deep fakes are precicely calibrated to feed back into and reinforce existing ideological biases, beliefs, and prejudices. This spawns further social divisions and political partisanship. As Aviv Ovadya states, "In the wrong hands, synthetic media could deepen divisions in society – and in government – as it becomes difficult to tell what's real and what isn't."[73] In recent hearings by the Committee on Intelligence, the U.S. House of Representative stated that "Deepfakes raise profound questions about national security and democratic governance, with

[70] Simon Parkin, "The Rise of the Deepfake and the Threat to Democracy," *The Guardian*, June 22, 2019, https://www.theguardian.com/technology/ng-interactive/2019/jun/22/the-rise-of-the-deepfake-and-the-threat-to-democracy.

[71] Aviv Ovadya, "Deepfake Myths: Common Misconceptions About Synthetic Media," *Alliance for Securing Democracy*, June 14, 2019, https://securingdemocracy.gmfus.org/deepfake-myths-common-misconceptions-about-synthetic-media/.

[72] Parkin, "The Rise of the Deepfake and the Threat to Democracy."

[73] Ovadya, "Deepfake Myths."

individuals and voters no longer able to trust their own eyes or ears when assessing the authenticity of what they see on their screens."[74] Deepfakes are a particularly insidious form of computational propaganda, and they have the potential to inflame tensions between nations, endanger national security, and undermine foreign policy and international diplomacy.[75] McBeth has even raised the possibility that an AI-engineered event could be so inflammatory as to launch an armed conflict.[76] In a political climate of partisanship, high tensions, and a failure of non-violent dispute-resolution mechanisms, this remains a possibility.

There is evidence that synthetic videos are beginning to be used by political opponents, although to date these fakes have been relatively easy to detect. In one such video, posted by the Flemish Socialist Party of Belgium, U.S. President Donald Trump appears to criticize Belgium for remaining in the Paris climate accord.[77] The Party told *Politico* that it created the video in order to start a public debate, and they were open about its falsity.[78] Another video surfaced in May of 2019 by a Trump supporter which depicted Nancy Pelosi, the Speaker of the U.S. House of Representatives, appearing to deliver a speech while highly intoxicated. Again, the video was quickly debunked,[79] but with media credibility at an all-time low, there will be many who prefer

[74] Permanent Select Committee on Intelligence, "House Intelligence Committee to Hold Open Hearing on Deepfakes and AI: The National Security Challenge of Artificial Intelligence, Manipulated Media, and 'Deepfakes,'" Press Release, U.S. House of Representatives Permanent Select Committee on Intelligence, Washington, D.C., June 7, 2019, https://intelligence.house.gov/news/documentsingle.aspx?DocumentID=657.

[75] Danielle Keats Citron, "Prepared Testimony and Statement for the Record of Danielle Keats Citron, Morton and Sophia Macht Professor of Law, University of Maryland Carey School of Law, Hearing on "The National Security Challenge of Artificial Intelligence, Manipulated Media, and 'Deepfakes'" Before the House Permanent Select Committee on Intelligence, June 13, 2019, https://intelligence.house.gov/uploadedfiles/citron_testimony_for_house_committee_on_deep_fakes.pdf, 6.

[76] Katie McBeth, "The Infocalypse is Coming," International Policy Digest, March 8, 2018, https://intpolicydigest.org/2018/03/08/the-infocalypse-is-coming/.

[77] sp.a, Facebook, May 19, 2018, https://www.facebook.com/Vlaamse.socialisten/videos/10155618434657151/.

[78] Hans Von Der Burchard, "Belgian Socialist Party Circulates 'Deep Fake' Donald Trump Video," *Politico*, May 21, 2018, https://www.politico.eu/article/spa-donald-trump-belgium-paris-climate-agreement-belgian-socialist-party-circulates-deep-fake-trump-video/.

[79] Simon Parkin, "The Rise of the Deepfake and the Threat to Democracy," *The Guardian*, June 22, 2019, https://www.theguardian.com/technology/ng-interactive/2019/jun/22/the-rise-of-the-deepfake-and-the-threat-to-democracy.

to believe in the truth of the video anyways – to 'debunk the debunking' – and entrench existing biases.

In 2018, deep fake photos were posted on conservative media sites depicting Parkland school shooting survivor and gun-control activist Emma Gonzalez.[80] In the doctored photos, Gonzalez was shown wearing a Cuban flag on her jacket and ripping up the U.S. Constitution. These photos inflamed conservatives and Second Amendment aficionados, and in this case the photos were *not* quickly debunked as fakes. In one the original photos, Gonzalez is actually ripping up a shooting target, not the U.S. Constitution; more than this, the original photo appeared prominently in *Teen Vogue* as part of a cover story on school shootings.[81] This shows just how deeply entrenched filter bubbles can be in such a vast and divided media landscape, and how easily even obvious synthetic media can go undetected and support political and social divisions.

At the same time, the phenomena will surely arise whereby any public figure caught in an embarrassing or compromising act will simply claim it to be a deep fake, and blame it on an algorithm – a phenomenon that Citron has called the Liar's Dividend, stating that the "more people are educated about the advent of deep fakes, the more they may disbelieve real recordings. Regrettably and perversely, the Liar's Dividend grows in strength as people learn more about the dangers of deep fakes."[82]

Synthetic media therefore have the capacity to destroy social capital and credibility at its base, and this is equally true whether we accept their content as true or not. The more

[80] Maura Barrrett and Jo Ling Kent, "Inside the Government Agency Designing Tech to Fight Fake News," NBC News, April 18, 2018, https://www.nbcnews.com/tech/tech-news/inside-government-agency-designing-tech-fight-fake-news-n865586.
[81] Barrett, Inside the Government Agency."
[82] Citron, "Prepared Testimony and Statement for the Record," 7.

fundamental problem is that we do not have any objective criteria to determine what is worthy of belief. Moreover, drawing attention to the phenomenon of synthetic media only reinforces its effects, leaving us no choice but to fall back on our preferences, prejudices, and political ideologies – exactly the opposite of what we need to reduce social discord and promote healthy civic engagement.

Shamir Alibhai is the CEO of Amber, a company that embeds a watermark in videos at the time of their creation to assist in their authentication. Alibhai believes that authenticating media in this way is a matter of deep moral principle. He states that "a postfact world could undo much of the last century's progress toward peace, stability and prosperity, driven in part by a belief in evidence-based conclusions."[83] Similarly, Aviv Ovadya states:

> More generally, synthetic media is a challenge to our *epistemic capacity* – our ability to make sense of the world and make competent decisions. Especially concerning is the growth of *reality apathy* – where people give up on determining real from fake – and *reality sharding* – where people selectively choose what to believe, forming deeper and deeper like-minded clusters.[84]

This kind of 'reality apathy,' or 'reality fatigue,' is a symptom of eroding epistemic capacity and the credibility of core civic institutions – the media, politicians, public figures, academia. "Beset by a torrent of constant misinformation," he states, "people simply start to give up."[85]

Reality apathy may lead to a fundamental loss in the credibility of social institutions, a crisis that Ovadya has termed the 'infocalypse'[86] – the point at which we realize that we lack standards for truth and accuracy, and that we have no control over fast-evolving technologies.

[83] Parkin, "The Rise of the Deepfake and the Threat to Democracy."
[84] Aviv Ovadya, "Deepfake Myths: Common Misconceptions About Synthetic Media," Alliance for Securing Democracy, June 14, 2019, https://securingdemocracy.gmfus.org/deepfake-myths-common-misconceptions-about-synthetic-media/.
[85] Charlie Warzel, "He Predicted the 2016 Fake News Crisis. Now He's Worried About an Information Apocalypse," *Buzzfeed News*, February 11, 2018, https://www.buzzfeednews.com/article/charliewarzel/the-terrifying-future-of-fake-news#.auEWOgLKA.
[86] Warzel, "He Predicted the 2016 Fake News Crisis."

Ovadya states that “the stakes are high and the possible consequences more disastrous than foreign meddling in an election – an undermining or upending of core civilizational institutions, an ‘infocalypse’.”[87]

4.3 The Search Engine Manipulation Effect (SEME):

Algorithmic injustice includes a variety of inequalities and deprivations that stem from the operation of algorithms and machine-learning systems. This can include, for example, a tendency to be mis-identified by a facial recognition system because of one’s race, and other biases that are inherent in machine-learning systems and the way that information is presented to them during their learning process.[88] Algorithmic injustices can promote bias and discrimination and can lead to differential access to goods and services. Since algorithms also control an increasing variety of security systems, they can also have an outsized impact on our privacy, on the security of our personal information, and our susceptibility to surveillance and data collection.

Moreover, many of these risks are invisible, buried as they are inside impenetrable and proprietary algorithms – the so-called ‘black box’ that is a fundamental characteristic of machine-learning systems. As Jonathan Zittrain states, “most machine-learning systems don’t uncover causal mechanisms. They are statistical-correlation engines.”[89] *They* can’t – and so *we* can’t – explain why they make the findings that they do.[90] As Epstein states, “Google decides which of the billions of web pages it is going to include in our search results, and it also decides

[87] Warzel, “He Predicted the 2016 Fake News Crisis.”
[88] Shoshana Amielle Magnet, *When Biometrics Fail: Gender, Race, and the Technology of Identity* (Durham: Duke University Press, 2011).
[89] Johnathan Zittrain, “The Hidden Costs of Automated Thinking,” The New Yorker, July 24, 2019, https://www.newyorker.com/tech/annals-of-technology/the-hidden-costs-of-automated-thinking.
[90] Johnathan Zittrain, “The Hidden Costs of Automated Thinking,” The New Yorker, July 24, 2019, https://www.newyorker.com/tech/annals-of-technology/the-hidden-costs-of-automated-thinking.

how to rank them. How it decides these things is a deep, dark secret."[91] Since these systems are largely independent from human knowledge and control, we cannot really know how or why they are operating the way they do, and so we cannot penetrate the black box to uncover the roots of systemic biases.

All of these problems hold out significant risks for our democratic values and equality rights. In this paper, we will focus on only one form of algorithmic injustice as it poses especial risks to the democratic process. This is the search engine manipulation effect (SEME), and its ability to invisibly shape public opinion and sway election votes. The SEME refers to the strong preference that people have for the information that appears at the top of the search rankings. Not only do we click far more often on the highest-ranked results, but we find the information therein to be far more credible and persuasive, even though we have minimal knowledge of how the results are ranked.[92] Google, either through its search engine or through its ownership of YouTube, controls what web sites and video content billions of people around the world see, and determines the timing of their access to this information.[93] For Google, over 90% of all clicks are on the first page of search results – with fewer than 10% of users venturing on to the second page to find a link – and about half of all clicks are on the first- or second-ranked link alone.[94] We have been conditioned to think of search engine results as being algorithmic, mathematical, and objective; this has led to powerful assumptions that the results at the top of the list are inherently trustworthy and more credible than those further down.[95]

[91] Robert Epstein, "The New Mind Control," *Aeon*, February 18, 2016, https://aeon.co/essays/how-the-internet-flips-elections-and-alters-our-thoughts.

[92] Robert Epstein and Ronald E. Robertson, "The Search Engine Manipulation Effect (SEME) and its Possible Impact on the Outcomes of Elections," PNAS 112, no. 33 (2015): E4512.

[93] Robert Epstein, "Why Google Poses a Serious Threat to Democracy, and How to End that Threat," Testimony Before the United States Senate Judiciary Subcommittee on the Constitution, June 16, 2019, 4.

[94] Epstein, "The Search Engine Manipulation Effect," E4512.

[95] Epstein, "Why Google Poses a Serious Threat to Democracy," 6.

SEME has been studied by Robert Epstein, Senior Research Psychologist at the American Institute for Behavioral Research and Technology.[96] He finds that the effect of biased search rankings on voter behavior is greater than that of traditional media sources,[97] and possesses an unprecedented ability to manipulate a significant portion of the world's population.[98] The problem is compounded by the virtual monopoly that Google has on search rankings, combined with the lack of transparency in their proprietary ranking methods serves to hide any bias or manipulation. Moreover, because Google can reach people at an unprecedented scale and at just the right time – such as during critical moments in election campaigns – this is a highly effective, and a wholly novel, form of social control.[99] As Epstein also points out, many elections are won by rather small margins, so even minimal manipulation can have an outsized influence on results provided it is targeted at undecided voters at just the right time, often in the days leading up to an election.[100]

One powerful form of election manipulation uncovered by Epstein is 'digital gerrymandering.'[101] This refers to the popular practice of campaigns getting people out to vote in the days leading up to an election. However, it is possible for a search engine to preferentially target these messages to supporters of only *one* campaign or candidate; because get-out-the-vote campaigns are so influential on election outcomes, even a small amount of targeted manipulation may be enough to swing a close election.[102]

It is also possible for a search engine to preferentially target undecided voters, and voters from particularly vulnerable groups. Using information collected from online profiles, the search

[96] Epstein, "Why Google Poses a Serious Threat to Democracy," 1.
[97] Epstein, "The Search Engine Manipulation Effect," E4512.
[98] Epstein, "Why Google Poses a Serious Threat to Democracy," 1.
[99] Epstein, "The Search Engine Manipulation Effect," E4513.
[100] Epstein, "The Search Engine Manipulation Effect," E4518.
[101] Epstein, "The Search Engine Manipulation Effect," E4518.
[102] Epstein, "The Search Engine Manipulation Effect," E4518.

engine can then target timely messages to these groups; such manipulation can be highly biased, yet undetectable by regulators.[103] As Epstein states, "restricting search ranking manipulations to voters who have been identified as undecided while also donating money to favored candidates would be an especially subtle, effective, and efficient way of wielding influence."[104] Google's autocomplete is also highly manipulative – these suggestions "can turn a 50/50 split among undecided voters into a 90/10 split," without people even being aware they are being manipulated.[105]

Yet another form of the SEME is referred to as the 'digital bandwagon' effect, which happens when one candidate or campaign gets pushed higher in the search engine rankings; this endows the candidate with a veneer of credibility and popularity that quickly overshadows and crowds-out other candidates and points of view.[106] This results in a feedback loop that continues to magnify the effect, described by Epstein as "the process by which search rankings affect voter preferences might interact synergistically with the process by which voter preferences affect search rankings, thus creating a sort of digital bandwagon effect that magnifies the potential impact of even minor search ranking manipulations."[107]

Not only is SEME powerful, but individuals appear to be particularly powerless to resist it – and this is true *even when people are informed that the effect is taking place and the search rankings they are viewing are biased*. As Epstein explains:

> SEME is one of the most powerful forms of influence ever discovered in the behavioral sciences, and it is especially dangerous because it is invisible to people – "subliminal," in effect. It leaves people thinking they have made up their own minds, which is very much an illusion. It also leaves no paper trail for authorities to trace. Worse still, the very few people who can detect bias in search results shift even farther in the direction

[103] Epstein, "The Search Engine Manipulation Effect," E4519.
[104] Epstein, "The Search Engine Manipulation Effect," E4520.
[105] Epstein, "Why Google Poses a Serious Threat to Democracy," 3.
[106] Epstein, "The Search Engine Manipulation Effect," E4519.
[107] Epstein, "The Search Engine Manipulation Effect," E4520.

> of the bias, so merely being able to see the bias doesn't protect you from it. Bottom line: biased search results can easily produce shifts in the opinions and voting preference of undecided voters by 20 percent or more – up to 80 percent in some demographic groups.[108]

Action against the SEME therefore must be calibrated to deal with its psychological effects, otherwise the measures taken will not work, or may even exacerbate the effect.

Epstein states that Google's control over information through search-engine rankings will continue to influence election results around the world:

> Robertson and I calculated that Google now has the power to flip upwards of *25 per cent of the national elections in the world* with no one knowing this is occurring. In fact, we estimate that, with or without deliberate planning on the part of company executives, Google's search rankings have been impacting elections for years, with growing impact each year. And because search rankings are ephemeral, they leave no paper trail, which gives the company complete deniability.[109]

Unchecked, then, search engine manipulation can impact democracy at its foundations – even if one's preferred party, position, or candidate prevails[110] – because it makes election results a consequence of manipulation by powerful actors and not the freely-expressed choices of voters. In fact, there is evidence that search-engine manipulations are already affecting people's choices, purchasing behaviours, habits, beliefs, and voting preferences worldwide, without a good deal of conscious awareness or reflection that this is occurring.[111]

5. Regulating Technological Threats to Civil Society

5.1 Introduction: Towards a New Regulatory Framework:

We have yet to devise effective methods to hold platforms accountable for computational

[108] Epstein, "Why Google Poses a Serious Threat to Democracy," 2.
[109] Epstein, "The New Mind Control."
[110] Epstein, "Why Google Poses a Serious Threat to Democracy," 4.
[111] Epstein, "The New Mind Control."

propaganda, or address algorithmic transparency.[112] Empowering states to control companies, and limit the user-generated content they post, is not necessarily an effective response. The danger exists that empowering the state to control platform companies will only increase the state's power to wield computational propaganda as a tool of economic, political, and foreign policy.[113] Regulatory solutions need to empower users – not organizations, companies, or states.[114]

To date, platform companies have responded to criticisms over the spread of misinformation by exerting more control over content on their platforms. Several platforms have introduced tools, for example, to control content, to filter news, to moderate user participation, and have put in place more and more mechanisms to demonetize or remove outright content that violates – what are often vague and opaque – terms of use. As Moore states:

> [T]his type of collaboration, of democratic states working closely with corporations that have such detailed knowledge of the minutiae of our political and social lives, raises a rather frightening prospect that neither George Orwell or Aldous Huxley fully imagined. A world in which governments have access to all our digital information and communication, and therefore almost complete knowledge of who we are, who we communicate with, and how we engage with politics – not only via their own systems but via those run by the information intermediaries. In addition to which, by outsourcing its means of surveillance and control there are few democratic mechanisms of transparency or accountability, with many citizens blithely unaware it is even happening[.][115]

If powerful interest assume control over platform companies and their content, whether this be governmental or corporate power, this has the potential to lead only to "the obscene power to

[112] Samantha Bradshaw, Lisa-Marie Neudert, and Philip N. Howard, "Government Responses to the Malicious Use of Social Media" Report, NATO Stratcom COE, Oxford Internet Institute, November 2018, https://comprop.oii.ox.ac.uk/wp-content/uploads/sites/93/2019/01/Nato-Report.pdf, 12.
[113] Woolley, "Introduction: Computational Propaganda Worldwide," 14.
[114] European Commission, "A Multi-Dimensional Approach to Disinformation," 25.
[115] Moore, "Tech Giants and Civic Power," 55.

decide what information humanity can see and how that information should be ordered."[116]

Platform companies are already able to wield power over content in ways that contradict their users' interests. Twitter, for example has allowed mainstream news producers to publish disturbing yet newsworthy footage, while it has banned ordinary users who wish to publish similar material.[117] Platform companies have also shown great deference to local laws, even illiberal laws that suppress speech. Twitter filters content by country, and abides by local laws restricting free speech.[118]

Since the 2016 U.S. election, over 40 countries around the world have passed or proposed new laws to deal with "fake news, social media abuse, and election interference."[119] Many of these laws empower tech companies to control content, and thereby also empower governments to suppress free speech and consolidate their own authoritarianism. The same is true for regulatory measures that seek to criminalize disinformation.[120] Several countries have now criminalized even the sharing of content vaguely termed "disinformation," including several authoritarian countries with poor human rights records and a lack of due process; these include Iran, Kuwait, Egypt, Indonesia, Malaysia, Russia, Saudi Arabia, and Tanzania.[121] Other countries have expanded their definitions of illegal content, and passed laws to strengthen their control over the media.[122]

At the same time, allowing platform companies to govern themselves runs the risk that

[116] Robert Epstein, "To Break Google's Monopoly on Search, Make its Index Public," *Bloomberg Businessweek*, July 15, 2019, https://www.bloomberg.com/news/articles/2019-07-15/to-break-google-s-monopoly-on-search-make-its-index-public.

[117] James Ball, "Twitter: From Free Speech Champion to Selective Censor," *The Guardian*, August 21, 2014, https://www.theguardian.com/technology/2014/aug/21/twitter-free-speech-champion-selective-censor.

[118] Dave Neal, "Twitter to Filter Content Geographically," *The Inquirer*, January 27, 2012, https://www.theinquirer.net/inquirer/news/2141809/twitter-filter-content-geographically.

[119] Bradshaw, "Government Responses to the Malicious Use of Social Media" 4.

[120] Bradshaw, "Government Responses to the Malicious Use of Social Media," 4.

[121] Bradshaw, "Government Responses to the Malicious Use of Social Media," 8.

[122] Bradshaw, "Government Responses to the Malicious Use of Social Media," 8-9.

they will simply introduce initiatives that impair public regulatory responses. As Gorwa states, “since 2016, platform companies have implemented multiple changes in response to public concern. These initiatives, which range from new advertising tools to changes as to how they interact with political campaigns, seem designed to head off possible avenues of regulation while also effectively maintaining the highly-profitable status quo.”[123]

Self-governance measures undertaken by companies may not be transparent, accessible, or fair to users. Klein states that criticisms of fake news have led many advertising companies to update their policies “to deny services to fake news publishers.”[124] In this way, accounts or revenue can be terminated with little warning or justification. This is already leading to a backlash. There may be lawsuits by those who are de-platformed and de-monetized, thereby forcing companies to go back and specify clear and justifiable criteria in their contractual terms.

As we argue above in the discussion of Microsoft and Backpage, we need not choose between the freedom of companies to distribute content on the one hand, and the power of states to regulate and suppress this content on the other. To do so only feeds back into the clash of empires. As Neudert states, “the debate on computational propaganda itself has become a highly politicized proxy war.”[125] The European Commission states in its recent report on combatting disinformation, “[a]ny form of censorship either public or private should clearly be avoided.”[126]

We can and should choose regulatory options that respect fundamental rights. Accordingly, neither governments nor platform companies should be allowed to interfere with editorial content.[127] Instead, we can implement methods of governance that increase resilience to

[123] Robert Gorwa, “What is Platform Governance?” *Information, Communication & Society* 22, no. 6 (2019: 862.
[124] Klein, “Fake News: A Legal Perspective,” 10.
[125] Neudert, “Germany: A Cautionary Tale,” 179.
[126] European Commission, “A Multi-Dimensional Approach to Disinformation,” 5.
[127] European Commission, “A Multi-Dimensional Approach to Disinformation,” 30.

disinformation,[128] and enhance transparency and accountability for the sources of information, its funding, and its credibility and reliability.[129] The human and civil rights laws that are already in place provide the basic legal and ethical framework for the future of platform governance, emphasizing the values of fairness, accountability, transparency, ethics. This is what shifts power and legal rights back to technology users, as both consumers and citizens.[130]

In the following section, we propose three organizing principles that can be used to direct regulatory responses to computational propaganda: 1) there should be informational transparency, including for algorithms and information flows; 2) power should be decentralized, and greater user control should be promoted; and 3) there should be a focus on efforts to improve media literacy, fact-checking and credibility. These principles are all linked, and together we propose that they are better able to deal with some of the specific harms of computational propaganda while also promoting democratic values, human rights, and civic engagement.

5.2 Transparency:

Transparency is a key principle in regulating platform companies and the spread of computational propaganda. Transparency is about enabling users to easily access relevant information, so that they can make their own determinations about the quality of the content, its credibility, and any sources of bias. Transparency also means algorithmic transparency, so that users can get information on how and why they are receiving certain material, or why information has been presented or recommended to them.[131] This also includes transparency over when bots or other technological means are being used to artificially inflate the apparent

[128] European Commission, “A Multi-Dimensional Approach to Disinformation,” 5.
[129] European Commission, “A Multi-Dimensional Approach to Disinformation,” 14.
[130] Gorwa, “What is Platform Governance,” 865.
[131] European Commission, “A Multi-Dimensional Approach to Disinformation,” 24.

popularity of a piece of information.[132] Focusing on transparency is a superior method of regulation than content controls, as it promotes the flow of relevant information, avoids censorship, and empowers users.

One example of regulation through transparency is the effort to inform users about the sources of funding for political advertising, and foreign funding of domestic political campaigns.[133] As Bradshaw states, "Some government initiatives focus on monitoring the information ecosystem and providing users with portals to report misinformation. At a regional level, The East StratCom Task Force provides monitoring, training, and capacity building for disinformation campaigns that affect European Union institutions and member state governments."[134] Italy, too, has begun a national monitoring initiative, and has "established a monitoring portal citizen can use to report instances of fake news for investigation in the run up to the next election."[135]

Another example of promoting transparency is alerting users to biases in search engine results, which has been shown to suppress the effects of the search-engine manipulation effect. It has been shown that about 8% of users will detect bias in search engine results without any warnings, but that about 25% of users will detect bias with a warning; this has been shown to be fairly consistent across experiments.[136]

Promoting transparency also involves opening up some of the impenetrable and proprietary algorithms and methods of data analyses that companies have to date kept hidden. As The European Commission states in its recent Report on disinformation:

Transparency of algorithms is also needed for integrity of elections and this cannot be

[132] European Commission, "A Multi-Dimensional Approach to Disinformation," 23.
[133] Bradshaw, "Government Responses to the Malicious Use of Social Media," 7.
[134] Bradshaw, "Government Responses to the Malicious Use of Social Media," 10.
[135] Bradshaw, "Government Responses to the Malicious Use of Social Media," 11.
[136] Epstein, "Suppressing the Search Engine Manipulation Effect," 12.

> conducted without data checking. This could be done by requiring platforms to use open application programming interface, that does not reveal the algorithm but its results. This would allow third parties to build software for data checking to monitor effects. Twitter, for instance, is making its open API available to many organizations and small research groups.[137]

In this way, transparency is closely linked to the next regulatory domain, which consists of methods that decentralize information and information control, and that put more power into the hands of users.

5.3 Decentralization & User Control:

One of the most important principles in a framework for regulating platform companies is to decentralize control out of the hands of large players – the companies themselves, the governments who have the power to regulate them, and even influential organizations and NGOs – and into the hands of end users. User control includes control over algorithms, recommendations and filtering, as well as over generating and accessing content, and information about potential biases about content.

The European Commission has put forth a number of proposals to increase user control:

> This may include the development of built-in tools/plug-ins and applications for browsers and smartphones, to empower users to better control access to digital information. In particular, platforms should consider ways to encourage users' control over the selection of the content to be displayed as results of a search and/or in news feeds. Such system should give to the user the opportunity to have content displayed according to quality signals. Moreover, content recommendation systems that expose different sources and different viewpoints around trending topics should be made available to users in online platforms.[138]

It can be seen that many of these controls are linked to measures of transparency we discussed in

[137] European Commission, "A Multi-Dimensional Approach to Disinformation," 24.

[138] European Commission, "A Multi-Dimensional Approach to Disinformation," 28.

the section above, showing how closely linked these domains are, and how beneficial changes in one domain can support and reinforce benefits in the other.

Co-governance, or shared governance, is another option that has been proposed. Gorwa recommends a co-governance model, which involves cooperation between technology companies and third-party organizations whose responsibility is to improve ethics and accountability:

> Civil society organizations, for example, have advocated for some kind of organization that could perform multiple functions ranging from investigating user complaints to creating ethical frameworks for platform companies, perhaps modelled after international press councils which set codes of conduct and standards for news organizations.[139]

One of the advantages of co-governance is that it can shift power to control content away from companies and towards civil society organizations. Such organizations may be given the power to identify and label content, to provide information about possible sources of bias, partisan funding, and credibility concerns. Independent third-party organizations might even be employed to assess complaints against platform providers.[140] The downsides of co-governance are that it empowers organizations over individuals and end-users. Although co-governance can be beneficial in some instances – such as promoting infrastructure for fact-checking, content-assessment, and complaints mechanisms – it might also work against individual interests. One such organization, NewsGuard, was being employed to assess the credibility of new sources but instead has its own biases and partisan sources of funding exposed by those same news organizations.[141] This shows the dangers of empowering third-party organizations, since their

[139] Gorwa, “What is Platform Governance,” 864.

[140] Gorwa, “What is Platform Governance,” 864-5.

[141] Whitney Webb, “How a NeoCon-Backed ‘Fact-Checker’ Plans to Wage War on Independent Media,” Mint News, January 9, 2019, https://www.mintpressnews.com/newsguardneocon-backed-fact-checker-plans-to-wage-war-on-independent-media/253687/.

own transparency, credibility, and biases, can be very difficult to assess.

Another, more radical, method of promoting decentralization and user control is to declare search engine indexes, certain databases, and even some algorithms to be part of the public commons. Epstein, for example, has proposed the break-up of Google's monopoly by declaring its search index to be a public commons. He states that we should declare:

> Google's massive search index – the database the company uses to generate search results – to be a public commons, accessible by all, just as a 1956 consent decree forced AT&T to share all its patents. There is precedent in both law and in Google's own business practices to justify taking this step.[142]

Google may have gathered the content, and developed ways to index and search it, but that content was generated by users, and they should have control over how it is used.[143]

Search indices like Google's could be shared with third parties through an application programming interface, or API, much as Google already does with Startpage, a Dutch company.[144] Moving control over at least some digital functions into the public sphere is a more effective way to promote competition among search indices than are anti-trust actions, such as those taken against Microsoft. This could serve to break-up the search engine market and promote competition. As Epstein states:

> Declaring Google's index a commons will quickly give rise to thousands of search platforms like Google.com, each competing with Google, each providing excellent search results, each serving niche audiences, large and small, exactly like newspapers and television networks and websites do now. Search will become competitive, as it was during its early years, and democracy will be protected from Google's secretive machinations.[145]

As with all competition, this could promote real alternatives to Google, as it would create

[142] Epstein, "Why Google Poses a Serious Threat to Democracy," 5.
[143] Epstein, "To Break Google's Monopoly on Search."
[144] Epstein, "To Break Google's Monopoly on Search."
[145] Epstein, "Why Google Poses a Serious Threat to Democracy," 5.

"thousands of search platforms, each with its special focus and emphasis, each drawing on different subsets of information from Google's ever-expanding index, and each using different rules to decide how to organize the search results they display."[146]

Search engines open to the public via an API can thereby give users control over their search preferences, and this promotes competition as well as user satisfaction. Startpage doesn't track users' online activities, and so does not give personalized results like Google does, but it better maintains users' privacy.[147] With competition, Google may have to service its customer's interests, and even reign in some of its unpopular forms of surveillance.[148] This is one reason why search engines like Google may wish to adopt these methods, before they lose more search traffic to privacy and security-conscious sites, such as Duck-Duck-Go. Epstein argues that it may even be in Google's best interests to devolve some control into the commons, as it would avoid the more damaging scenario of having its assets seized, fined, or frozen by a government.[149]

Devolving some of the assets of technology companies into the commons is a regulatory option in Europe. Under European competition law, and its 'essential facility doctrine,' a dominant firm may be required to share its assets if the asset is essential or indispensable for others to compete effectively in the marketplace, and if refusing this access "would eliminate effective competition on that market and thus cause consumer harm."[150] There is also some precedent for this, as the European Commission has previously held that "Google has abused its market dominance as a search engine by giving an illegal advantage to another Google product,

[146] Epstein, "To Break Google's Monopoly on Search."
[147] Epstein, "To Break Google's Monopoly on Search."
[148] Nathan Newman, "Why Google's Spying on User Data is Worse than the NSA's," *Huffington Post*, November 20, 2015, https://www.huffpost.com/entry/why-googles-spying-on-use_b_3530296.
[149] Epstein, "To Break Google's Monopoly on Search."
[150] European Parliament, "Parliamentary Questions: Answer Given by Ms. Vestager on Behalf of the European Commission," Question Reference E-000408/2019, May 2, 2019, http://www.europarl.europa.eu/doceo/document/E-8-2019-000408-ASW_EN.html.

its comparison shopping service."[151] This was found to be a breach of EU antitrust rules.

5.4 Enncing Media Literacy & Credibility:

Educating users comprises a number of related matters – not only on media literacy, the standards of good journalism, and assessing sources of media bias, but also in promoting independent fact-checking organizations, and measures to assess and enhance media credibility. Efforts to educate and promote media literacy and credibility work together with the other two domains of regulation, as they promote user control as well as the transparency of information, its sources, and its funding.

Media literacy is an important component of regulating platform companies, but it must be specifically designed to address computational propaganda. Many technological forms of manipulation are psychologically powerful because they work in concert with our existing psychological biases. Education will be most effective if it specifically targets these effects. Attaching bias or other warnings to content may not be effective in and of itself. Pennycook *et al.* have shown that it may even have the opposite effect, such that "attaching warnings to headlines that have been flagged by third-party fact-checkers as disputed or unreliable in some way can have a negative effect such that readers see untagged headlines as somehow being more reliable, or having been validated."[152] The authors term this the "implied truth effect."[153] Attaching quality signals to information may only have a beneficial effect on users' perceptions if unflagged and unverified materials are also identified as such.

[151] European Commission, "Antitrust: Commission Fines Google €2.42 Billion for Abusing Dominance as Search Engine by Giving Illegal Advantage to Own Comparison Shopping Service" Press Release, June 27, 2017, https://europa.eu/rapid/press-release_IP-17-1784_en.htm.

[152] Gordon Pennycook, Adam Bear, Evan Collins, and David G. Rand, "The Implied Truth Effect: Attaching Warnings to a Subset of Fake News Headlines Increases Perceived Accuracy of Headlines Without Warnings," Management Science, Forthcoming (August, 2019): http://dx.doi.org/10.2139/ssrn.3035384.

[153] Pennycook, "The Implied Truth Effect."

The European Commission has also suggested that personalized content-filtering be programmed so as to facilitate "the display online of the widest possible range of relevant news and information to consumers."[154] Algorithms that recommend content should ensure that users are exposed to a greater, rather than a lesser, variety of content and viewpoints. This may address the growth and entrenchment of online filter bubbles. To ensure that the recommended sources are themselves not biased, they can be generated by a randomizing function. To ensure that this doesn't violate the principle of facilitating user control, users should always be able to opt out in their preferences.

Aviv Ovadya, through his Thoughtful Technology Project, has put together a framework that seeks to build the credibility of information sources. Credibility in this context means more than just accuracy in terms of factual content, but refers more broadly to information that improves our accurate understanding of the world and provides a basis for making sound decisions; misleading information, in Ovadya's framework, is that which *decreases* the accuracy of our beliefs about the world and the effectiveness of our decisions.[155] Efforts to promote credibility include increased media literacy, research, and funding for fact-checkers and other media verification specialists who, unlike NewsGuard, are independent and trustworthy.[156] Broad-based efforts at media literacy should begin from a young age.[157] This also means training for teachers and resources for curriculum development,[158] and for parents and libraries, as well. Europe needs to fund research, fact-checking, media literacy, as well as funding independent news media and training journalists, and de-emphasize government initiatives.[159] "The final

[154] European Commission, "A Multi-Dimensional Approach to Disinformation," 27.
[155] Aviv Ovadya, "What is Credibility Made Of?" Tow Center for Digital Journalism, March 21, 2019, https://www.cjr.org/tow_center_reports/ovadya-credibility-journalism-ocasio.php.
[156] European Commission, "A Multi-Dimensional Approach to Disinformation," 22.
[157] European Commission, "A Multi-Dimensional Approach to Disinformation," 26.
[158] European Commission, "A Multi-Dimensional Approach to Disinformation," 27.
[159] European Commission, "A Multi-Dimensional Approach to Disinformation," 28.

goal," states Ovadya, "should be the creation of an open market for fact-checking that avoids a 'monopoly of truth,' which could be potentially abused in some countries and might not carry public approval in other countries."[160]

6. Conclusion:

As we write, the U.S. Defense Advanced Research Projects Agency (DARPA) has put out a call for potential partners to develop their Semantic Forensics Program (SemaFor).[161] SemaFor will develop semantic detection algorithms to identity misinformation or manipulated and synthetic media, such as text, news, photos, and videos; attribution algorithms will identify the organizations and individuals who are the sources of the information, and other algorithms will stop the spread of the information for the purposes of helping to "identify, understand, and deter adversary disinformation campaigns."[162] This is a good example of an approach to media regulation that is not in keeping with the framework we proposed. First, it uses technological means, namely code, to solve what is essentially a problem involving social relations, critical thinking, and political values. Second, it proposes content controls to direct and limit the transmission of even lawful expressions. Third, it violates the core regulatory goals we have described here: it is not transparent – its algorithms will operate invisibly and are able to identify and remove content without users being aware; it takes control over information away from individuals, and centralizes it in the hands of the U.S. military; it spreads fear about the dangerousness of the media and information from vaguely-defined adversaries, sows distrust in

[160] European Commission, "A Multi-Dimensional Approach to Disinformation," 24.
[161] Defence Advanced Research Projects Agency (DARPA) (U.S.), "Semantic Forensics (SemaFor) Proposers Day (Archived)," DARPA, August 28, 2019, https://www.darpa.mil/news-events/semantic-forensics-proposers-day. This program is a more advanced iteration of the Media Forensic (MediFor) program discussed above, in Barrett, "Inside the Government Agency."
[162] DARPA, "Semantic Forensics."

the government for secretly removing access to lawful content, and it forces people to go outside of the mainstream to access news and information – this only encourages filter bubbles and access to misinformation and fake news.

We have proposed here a regulatory framework that is intended to maintain human rights and freedom of speech, promote competition and decentralization, and devolve power and control onto individuals as consumers, citizens, and technology users. This approach to Internet governance emphasizes goals, desired outcomes, and civic values – including political autonomy, open debate and freedom of expression. At the same time, this framework gives us governance mechanisms that are more effective in responding to the threats posed by the power of technology companies and the spread of computational propaganda.

In the 13th century, the leaders of the Abbasid empire, in their hubris, did not think that they were vulnerable to a Mongol invasion; they failed to respond adequately to the threat, and their empire – a cultural triumph centuries old – fell quickly. We should not allow our own governments to take ineffective measures against computational propaganda and the monopolistic power of technology companies. Nor should we allow our governments, in their hubris, to think that they can control the tech companies and the public narrative in order to shore up their own power. They can't. But in attempting to do so they *can* limit our basic freedoms, exacerbate social discord, and drive an ever-increasing number of people to seek out information sources farther from the mainstream – sources that lack credibility and that promote misinformation and conspiracy theories. The problem at hand is not one of technology, and it can only be solved by building trust in media sources and social institutions, ensuring that their credibility is deserved, and by promoting political autonomy, critical thinking, and the free flow of information. The Internet is an unprecedented vehicle for promoting education, spreading knowledge, and facilitating social relationships and political associations. This outcome, however, is not inevitable – it depends on the approach we take to governing the Internet and the spread of information. If we allow the Internet to be caught up in a monumental clash of empires, it, too, can come crashing down in a heartbeat.

* * *