September 2018

Platforms, the First Amendment and Online Speech: Regulating the Filters

Sofia Grafanaki

Follow this and additional works at: https://digitalcommons.pace.edu/plr

Part of the Computer Law Commons, First Amendment Commons, and the Internet Law Commons

Platforms, the First Amendment and Online Speech: Regulating the Filters

Sofia Grafanaki[*]

Abstract

In recent years, online platforms have given rise to multiple discussions about what their role is, what their role should be, and whether they should be regulated. The complex nature of these private entities makes it very challenging to place them in a single descriptive category with existing rules. In today's information environment, social media platforms have become a platform press by providing hosting as well as navigation and delivery of public expression, much of which is done through machine learning algorithms. This article argues that there is a subset of algorithms that social media platforms use to filter public expression, which can be regulated without constitutional objections. A distinction is drawn between algorithms that curate speech for hosting purposes and those that curate for navigation purposes, and it is argued that content navigation algorithms, because of their function, deserve separate constitutional treatment. By analyzing the platforms' functions independently from one another, this paper constructs a doctrinal and normative framework that can be used to navigate some of the complexity.

The First Amendment makes it problematic to interfere with how platforms decide what to host because algorithms that implement content moderation policies perform functions analogous to an editorial role when deciding whether content should be censored or allowed on the platform. Content navigation algorithms, on the other hand, do not face the same doctrinal challenges; they operate outside of the public discourse

[*] Sofia Grafanaki researches the personal and social implications of new technologies. She focuses on the desirability and possibility of their democratic supervision. The author has worked extensively with startups in the Big Data space, holds two L.L.M. degrees from NYU School of Law, an M.B.A. from Columbia Business School, and a law degree from Oxford University. She is admitted to practice law in New York, London, and Athens.

111

as mere information conduits and are thus not subject to core First Amendment doctrine. Their function is to facilitate the flow of information to an audience, which in turn participates in public discourse; if they have any constitutional status, it is derived from the value they provide to their audience as a delivery mechanism of information.

This article asserts that we should regulate content navigation algorithms to an extent. They undermine the notion of autonomous choice in the selection and consumption of content, and their role in today's information environment is not aligned with a functioning marketplace of ideas and the prerequisites for citizens in a democratic society to perform their civic duties. The paper concludes that any regulation directed to content navigation algorithms should be subject to a lower standard of scrutiny, similar to the standard for commercial speech.

Introduction

In February 2018, Facebook and Instagram were singled out by the Justice Department, more than any other technological tool, as critical to Russian efforts to influence the 2016

Presidential Election.[1] Facebook handed over to Congress more than 3,000 ads that were purchased by the Russian troll farm known as the Internet Research Agency,[2] which, according to the company, reached eleven million of its users.[3] Research however indicates, that the organic reach of the Russian-controlled accounts was dramatically bigger; Facebook users may have been exposed to such content hundreds of millions, or perhaps billions, of times.[4] Shortly after the election, a news analysis found that during the last few months of the US Presidential campaign, the top-performing fake news stories on Facebook generated more engagement than the top stories from major news outlets.[5] On April 10th and 11th of 2018, Facebook CEO

---

1. Sheera Frenkel & Katie Benner, *To Stir Discord in 2016, Russians Turned Most Often to Facebook*, N.Y. TIMES (Feb. 17, 2018), https://www.nytimes.com/2018/02/17/technology/indictment-russian-tech-facebook.html.

2. *See generally* Adrian Chen, *The Agency*, N.Y. TIMES MAG., June 2, 2015, at 57; Neil MacFarquhar, *Inside the Russian Troll Factory: Zombies and a Breakneck Pace*, N.Y. TIMES, Feb. 19, 2018, at A11; Mike Isaac & Scott Shane, *Facebook to Deliver 3,000 Russia-Linked Ads to Congress on Monday*, N.Y. TIMES (Oct. 1, 2017), https://www.nytimes.com/2017/10/01/technology/facebook-russia-ads.html; *see also* Dylan Byers, *Facebook Handed Russia-Linked Ads Over to Mueller Under Search Warrant*, CNN BUS. (Sept. 17, 2017, 11:29 AM), https://money.cnn.com/2017/09/15/media/facebook-mueller-ads/index.html?sr=fbCNN091517facebook-mueller-ads0907PMStoryLink&fbclid=IwAR3C0D5OQteU9ogfkny_loixjJNDCrRhToQc2DRwYCrcEMcBxlTHo2rm4p4; Benjamin Siegel, *Facebook Turns Over Thousands of Russia-Linked Ads to Congress*, ABC NEWS (Oct 2, 2017, 5:04 PM), https://abcnews.go.com/Politics/facebook-turns-thousands-russia-linked-ads-congress/story?id=50226525.

3. *Facebook, Social Media Privacy, and the Use and Abuse of Data: Hearing Before the U.S. S. Comm. on the Judiciary and the U.S. S. Comm. on Commerce, Sci. and Transp.*, 115th Cong. 5 (2018) (testimony of Mark Zuckerberg, Chairman and Chief Executive Officer, Facebook), https://www.judiciary.senate.gov/imo/media/doc/04-10-18%20Zuckerberg%20Testimony.pdf.

4. Craig Timberg, *Russian Propaganda May Have Been Shared Hundreds of Millions of Times, New Research Says*, WASH. POST (Oct. 5, 2017), https://www.washingtonpost.com/news/the-switch/wp/2017/10/05/russian-propaganda-may-have-been-shared-hundreds-of-millions-of-times-new-research-says/?utm_term=.ffb82f4fe621; *see generally Itemized Posts and Historical Engagement - 6 Now-Closed FB Page*, TABLEAU PUB., https://public.tableau.com/profile/d1gi#!/vizhome/FB4/TotalReachbyPag (last updated Oct. 5, 2017).

5. Craig Silverman, *This Analysis Shows How Viral Fake Election News Stories Outperformed Real News on Facebook*, BUZZFEED NEWS (Nov. 16, 2016, 5:15 PM), https://www.buzzfeed.com/craigsilverman/viral-fake-election-news-outperformed-real-news-on-facebook?utm_term=.yxER8laNQ#.avAbrgen7.

Mark Zuckerberg testified before the Senate and two Congressional Committees after a scandal broke out that political consulting firm, Cambridge Analytica, had improperly obtained personal information on eighty-seven million Facebook users.[6] The CEOs of Google and Twitter were also called to testify.[7] The Facebook hearings were supposed to be about Facebook's data privacy practices, but questions came up on a wide range of topics, such as the existence of political bias within the company, Facebook's role in the opioid crisis, whether Facebook is responsible for the content on its platform the same way publishers are, and whether Facebook is a monopoly.[8] Cambridge Analytica subsequently announced that it was shutting down.[9]

These incidents represent only the latest wave of issues involving online platforms. In recent years, the platforms' many functions have given rise to disagreements both in academic literature and pop culture about what their role is and what it should and should not be. Are platforms media companies or technology companies? Are they broadcasters or mere intermediaries? Are they editors with First Amendment protection, or do they behave like state actors? Do they operate like public utilities? Are they monopolies?[10] The list of questions continues, but underlying the growing body of literature is a

---

6. *See* Nicholas Confessore, *Cambridge Analytica and Facebook: The Scandal and the Fallout So Far*, N.Y. TIMES (Apr. 4, 2018), https://www.ny times.com/2018/04/04/us/politics/cambridge-analytica-scandal-fallout.html.

7. Alfred Ng, *Senate Summons Facebook, Google, Twitter CEOs over Data Privacy*, CNET (Mar. 26, 2018, 1:23 PM), https://www.cnet.com/ news/senate-calls-google-facebook-twitter-ceos-in-for-data-privacy/.

8. *See Transcript of Mark Zuckerberg's Senate Hearing*, WASH. POST (Apr. 10, 2018), https://www.washingtonpost.com/news/the-switch/wp/2018/04 /10/transcript-of-mark-zuckerbergs-senate-hearing/?utm_term=.e87c9e091a48; *Transcript of Zuckerberg's Appearance Before House Committee*, WASH. POST (Apr. 11, 2018), https://www.wash ingtonpost.com/news/the-switch/wp/2018/04/11/transcript-of-zuckerbergs-appearance-before-house-committee/?utm_term=.cec873708c75.

9. Rebecca Ballhaus & Jenny Gross, *Cambridge Analytica Closing Operations Following Facebook Data Controversy*, WALL ST. J. (May 2, 2018, 7:20 PM), https://www.wsj.com/articles/cambridge-analytica-closing-opera tions-following-facebook-data-controversy-1525284140.

10. For an analysis of the different analogies used, *see generally* Heather Whitney, *Search Engines, Social Media, and the Editorial Analogy*, KNIGHT FIRST AMENDMENT INST. (Feb. 26, 2018), https://knightcolumbia.org /content/search-engines-social-media-and-editorial-analogy.

common set of concerns: should we regulate the way these platforms handle public expression, and, if so, can we do it in a way that is aligned with our constitutional values?

This paper answers both concerns in the affirmative, arguing that there is a subset of algorithms social media platforms use to filter public expression, which can be regulated without constitutional barriers. These are the *content navigation algorithms*,[11] which, because of their function, deserve separate constitutional treatment. Unlike algorithms that decide what content to censor or allow on the platform, which perform a somewhat editorial role and, as such, receive constitutional protection, the regulation of content navigation algorithms is not preempted by the First Amendment. This categorization is elaborated on below, but to make the point less abstract, an application is offered: while we cannot address a phenomenon like *fake news* by outlawing its existence or requiring platforms to censor it, what we *can* do is address the way fake news gets *amplified* by regulating the algorithms that perform content navigation functions. Fake news is only one of the issues that such regulation can address; it serves as a timely example because it is an issue that, perhaps better than any other, has highlighted the pathologies of the current digital information ecosystem.[12]

A first step to unpacking the issues is to clarify what we mean when we refer to online *platforms*. Social media companies were the first to use the term *platform* in describing their services, but the term now extends to companies focused on services beyond expressive content, such as retail (Amazon), transportation (Uber), and real estate and hospitality (AirBnB).[13] For the purposes of this paper, the analysis is limited to the original use of the term, i.e. to social media platforms that handle predominantly expressive content. Examples include: Facebook, Instagram, Twitter, and YouTube.

---

11. *See infra* Part II(B) for definition and discussion of content navigation algorithms.

12. *See generally* David M. J. Lazer et al., *The Science of Fake News*, 359 SCI. 1094 (2018), http://science.sciencemag.org/content/359/6380/1094.

13. Tarleton Gillespie, *The Platform Metaphor, Revisited*, CULTURE DIGITALLY (Aug. 24, 2017), http://culturedigitally.org/2017/08/platform-metaphor/.

116   *PACE LAW REVIEW*   Vol. 39.1

Retail and other platforms operate under different economics and should be analyzed separately. Google, although it has somewhat similar economics to social media platforms, is also distinguishable in the way it curates public expression.

Social media platforms are private entities that play two distinct roles in the systems of information flow: 1) they host public expression online, i.e. they offer storage, and 2) they provide "navigation and delivery of the digital content of others."[14] They operate as online content intermediaries, whose functions are performed by algorithms that are designed by humans, with humans sometimes interfering with the results of the algorithms.

Interestingly, commentators have pointed out that the choice of the word *platform* is not accidental, both in the companies' self-characterizations and in the public discourse. Rather, the term reveals the position that such intermediaries are trying to establish.[15] A platform connotes "a 'raised, level surface' designed to facilitate some activity that will subsequently take place."[16] The term is "anticipatory, but not causal," implies an initial neutrality, and suggests a progressive and egalitarian arrangement that promises to "support those who stand upon it."[17]

As such, the word platform seems to serve social media companies quite well. Social media companies have to appeal to different constituencies, namely users, advertisers, content producers, and policymakers, in a way that eases the tensions between them and implies equality and fairness.[18] Using such positioning, these private entities have become the primary curators of the cultural discussion online,[19] but have managed to

---

14. Tarleton Gillespie, *The Politics of 'Platforms,'* 12 NEW MEDIA & SOC'Y 347, 348 (2010); *see also* EMILY BELL & TAYLOR OWEN, TOW CTR. FOR DIG. JOURNALISM, THE PLATFORM PRESS HOW SILICON VALLEY REENGINEERED JOURNALISM (2017), http://towcenter.org/wp-content/uploads/2017/04/The_Platform_Press_Tow_Report_2017.pdf.

15. Gillespie, *id.* at 348.

16. *Id.* at 350.

17. *Id.*

18. *Id.* at 348.

19. *See* Elisa Shearer & Jeffrey Gottfried, *News Use Across Social Media Platforms 2017*, PEW RES. CTR. (Sept. 7, 2017), http://www.journalism.org/2017/09/07/news-use-across-social-media-platforms-2017/ (reporting that, according to a 2017 survey by the Pew Research Center, 67% of Americans get

present themselves as the antidote to traditional mass media associated with the notion of "elitist gatekeeper[s]."[20] Part I will elaborate on these dynamics, but for introductory purposes two (perhaps self-evident) observations are made. Firstly, the interests of the different constituencies involved are not aligned most of the time. Secondly, it is the economic interests of the platforms themselves as private companies that ultimately drive their business models.

Curating speech or expressive content can come in flavors, and there are several problematic issues when it comes to how these private entities curate speech. To make sense of them, this Article starts by making an important distinction. There are two different ways in which these private entities curate speech, which parallel their separate functions of *hosting* versus providing *navigation and delivery* of public expression.

The first way platforms curate or govern online speech is through their content moderation policies. The process of content moderation determines whether specific content items can be hosted and can continue to be hosted on the platform. It addresses the question of whether content can exist or survive on the platform (*ex ante* and *ex post* moderation[21]), regardless of who actually sees or interacts with it. This process involves monitoring for illegal, offensive, and inappropriate content, or content that is simply not consistent with the culture of the users, and subsequently deciding whether to censor such content.[22] Content moderation policies consist of detailed sets of rules, which are constantly updated and largely opaque. Their opaqueness has been long criticized by scholars[23] and has occasionally been the cause of public outcry in instances where

---

news from social media).

20.   Gillespie, *supra* note 14, at 352.

21.   *See id.* at 1636–38.

22.   *See* Kate Klonick, Comment, *The New Governors: The People, Rules, and Processes Governing Online Speech*, 131 HARV. L. REV. 1598 (2018) (arguing that "these platforms are best thought of as self-regulating private entities, governing speech within the coverage of the First Amendment by reflecting the democratic culture and norms of their users" (footnotes omitted)).

23.   *See generally* Danielle Keats Citron & Benjamin Wittes, *The Internet Will Not Break: Denying Bad Samaritans § 230 Immunity*, 86 FORDHAM L. REV. 401 (2017); Rebecca Tushnet, *Power Without Responsibility: Intermediaries and the First Amendment*, 76 GEO. WASH. L. REV. 986 (2008).

content is perceived as unjustly or wrongfully removed.[24] While concerns relating to content censorship decisions are very legitimate and thoroughly discussed by academics, they are not the focus of this paper.

The second way in which these platforms govern speech is by providing navigation through the infinitely growing quantity of available content. In this case, the curation process is not about allowing or disallowing content, but rather curation is about deciding to which specific content items to point the users' attention. It is this latter type of speech curation that is addressed in this paper.[25] Such curation is addressed both to users as individuals through personalization algorithms, and to all platform users as a whole, through algorithms that select content that is trending or popular. In both of these forms, content navigation algorithms are, to an increasing degree, shaping our participation in public life.[26] As Tarleton Gillespie has observed, "[t]ogether, these algorithms not only help us find information, they also provide a means to know what there is to know and how to know it, to participate in social and political discourse, and to familiarize ourselves with the publics in which we participate."[27]

The concerns in this type of speech curation are not about censorship *per se*. Algorithmic measures of popularity or what is trending make an explicit claim of a "calculated public[]" and become cultural objects of meaning, in which individuals can look for "a reflection of the public in which they take part."[28]

---

24. *See* Zoe Kleinman, *Fury over Facebook 'Napalm Girl' Censorship*, BBC NEWS (Sept. 9, 2016), http://www.bbc.com/news/technology-37318031.

25. The two categories may seem to overlap sometimes, such as when a trending algorithm selects content that is considered illegal or inappropriate, but the questions of why the specific content was allowed to exist on the platform and why it surfaced as popular or trending are still fundamentally different questions.

26. Tarleton Gillespie, *The Relevance of Algorithms*, *in* MEDIA TECHS.: ESSAYS ON COMMC'N, MATERIALITY, & SOC'Y 167, 167 (Tarleton Gillespie et al. eds., MIT Press 2014).

27. *Id.*

28. Tarleton Gillespie*, #TrendingisTrending: When Algorithms Become Culture*, *in* ALGORITHMIC CULTURES: ESSAYS ON MEANING, PERFORMANCE & NEW TECHS. 52, 67, 69 (Robert Seyfert & Jonathan Roberge eds., 2016) (explaining that "calculated publics. . .imply a body of people who have been measured and assessed, as an explanation for why particular information has been presented as relevant to them" and that "trending algorithms make the claim of this

However, the ways in which this *public* is measured and shaped depart significantly from traditional assumptions on the role of news media in society and the democratizing promise of the Internet. This raises growing concerns about the failure of media literacy and lack of data literacy.[29] Personalization further adds to the concerns, as the fracturing of individual experiences, has been criticized for producing *filter bubbles* and *echo chambers*.[30] Users are only directed to content that is in agreement with their existing viewpoints, which in turn leads to ideological isolation, polarization in society, and increased vulnerability to believing in falsehoods—or so the argument goes.[31]

The distinction between the two types of speech curation just described—curating for *hosting* purposes versus curating for *navigation* purposes—matters especially when assessing the viability of legal regulation, and can help categorize the growing body of literature. Doctrinal barriers[32] make it extremely challenging to regulate how platform content moderation policies work, meaning how platforms make decisions about what content items to host. On the other hand, this paper will argue, the challenges are not as strong when it comes to how platforms facilitate content navigation. The assertion is that we *should* regulate content navigation algorithms to an extent, that

---

calculated public more explicit: this is what 'we' are reading, this is what my city or country is tweeting about, this is what America is listening to today.").

29. *See* Monica Bulger & Patrick Davison, *The Promises, Challenges, and Futures of Media Literacy* DATA & SOC'Y RESEARCH INST. 15–17 (2018), https://datasociety.net/pubs/oh/DataAndSociety_Media_Literacy_2018.pdf.

30. *See generally* ELI PARISER, THE FILTER BUBBLE: WHAT THE INTERNET IS HIDING FROM YOU (Penguin Books Ltd. 2011); CASS R. SUNSTEIN, #REPUBLIC: DIVIDED DEMOCRACY IN THE AGE OF SOCIAL MEDIA (Princeton Univ. Press 2017).

31. *See* Michela Del Vicario et al., *The Spreading of Misinformation Online,* 113 PROC. NAT'L ACAD. SCIS. U.S. AM. 554, 558 (2016), http://www.pnas.org/content/pnas/113/3/554.full.pdf (stating "[u]sers tend to aggregate in communities of interest, which causes reinforcement and fosters confirmation bias, segregation, and polarization. This comes at the expense of the quality of the information and leads to proliferation of biased narratives fomented by unsubstantiated rumors, mistrust, and paranoia"); *but see* Yochai Benkler et al., *Study: Breitbart-Led Right-Wing Media Ecosystem Altered Broader Media Agenda*, COLUMBIA JOURNALISM REVIEW (Mar. 3, 2017), https://www.cjr.org/analysis/breitbart-media-trump-harvard-study.php (suggesting that polarization is asymmetric).

32. Namely, the First Amendment combined with Section 230 of the Communications Decency Act.

we can do so without doctrinal barriers, and that doing so can address some of the challenges present in the content moderation context.

The paper proceeds in the following manner: Part I addresses the question of why regulate at all and explains what makes online platforms deserving of independent analysis. Part II addresses doctrinal barriers to regulation and argues that these do not apply in the content navigation context. Part III touches on an additional way to regulate the content navigation architecture and provides a brief supplementary discussion on the failure of self-regulation. Part IV concludes.

I. Why Regulate?

This part of the paper illustrates what is different in today's information ecosystem, why online platforms deserve independent analysis and why we should regulate at all.

A. Is There Anything Different Here?

Before embarking on the analysis of platforms, it is worth taking a brief pause and looking at a bigger picture. Back in 1996, at a conference on the *Law of Cyberspace*, Judge Frank Easterbrook flatly told the assembled crowd that no one in the room was going to win a Nobel Prize and that they were all at risk of multidisciplinary dilettantism.[33] For Judge Easterbrook, the *Law of Cyberspace* was as useless as the *Law of the Horse*.[34] His reasoning was that "the best way to learn the law applicable to specialized endeavors is to study the general rules."[35] Cases on people kicked by horses are better understood through the law of torts, not by reading 100% of the cases on this very specific issue.

Similarly, one could argue that, in some ways, there is nothing special about online platforms. We have the general rules on issues such as defamation, discrimination, competition,

---

33. Frank H. Easterbrook, *Cyberspace and the Law of the Horse,* 1996 U. CHI. LEGAL F. 207, 207 (1996).

34. *See id.*

35. *Id.*

free speech, intermediary liability, and so forth. This will not be the first time the law has to adapt to a new technology, nor is it the first time that companies driven by advertising profits select the headlines. Media companies have been doing this for years.[36] However, there *is* something very special about cyberspace and there *is* value in looking at these platforms separately. Presented below are two responses advanced by scholars that illustrate why in the context of cyberspace the issues deserve independent study.

　　Writing in 2004, Jack Balkin cautioned that "[i]n studying the Internet, to ask 'What is genuinely new here?' is to ask the wrong question."[37] For Balkin, focusing on novelty is the wrong way to think about technological change and public policy.[38] We will always find an analogue in the past, and then conclude that because there is nothing utterly new, nothing important has changed.[39] Instead, what we should be focusing on is *salience*. In Balkin's view, the right questions to ask are: which elements of the social world and what features of human activity or the human condition are emphasized and brought to the foreground by the new technology?[40] More importantly, what are the consequences of this new emphasis for human freedom?[41] In other words, even if there is nothing utterly novel, and the change is not a change in kind but simply one of degree, it can still have important consequences for society.

　　From a different point of view, Lawrence Lessig saw a structural change taking place in cyberspace, one that can teach very important general principles.[42] Lessig points out that law

---

　　36.　*See infra* note 56 and accompanying text.

　　37.　Jack M. Balkin, *Digital Speech and Democratic Culture: A Theory of Freedom of Expression for the Information Society*, 79 N.Y.U. L. REV. 1, 2 (2004).

　　38.　*Id.* at 2–3.

　　39.　*Id.*

　　40.　*Id.*

　　41.　*Id.*

　　42.　A few years post-Judge Easterbrook's intervention, Lawrence Lessig, who could not put the issue behind, wrote a commentary in the *Harvard Law Review*, as well as a more complete book, in an effort to illustrate that the Law of Cyberspace can, in fact, teach some very important general principles. *See generally* Lawrence Lessig, *The Law of the Horse: What Cyberlaw Might Teach,* 113 HARV. L. REV. 501 (1999); *see also* LAWRENCE LESSIG, CODE AND OTHER LAWS OF CYBERSPACE (Basic Books 1999).

is just one of the ways that human behavior can be regulated; there are three additional modalities of regulation, namely social norms, markets, and *architecture*.[43] These modalities regulate together, and any policy's ultimate effect should be seen as the sum of the regulatory effects of all four modalities together. Why does this matter for our purposes? The answer is that, in the digital environment (i.e. cyberspace), the fourth modality is quite different than it is in real space.

By *architecture*, Lessig refers to the physical world—both "as we find it" and "how it has already been made."[44] He gives the example of the city of Paris, where large boulevards limit the ability of revolutionaries to protest;—a constraint on behavior that has nothing to do with *legal* regulation. Similarly, Long Island bridges were built low, blocking public buses from passing under. The result was that people depending on public transport could not get to the beaches in Long Island, but the constraint was one imposed by architecture.

The non-plasticity of real space means there is very little we can do to change the architecture once it is there, or that doing so is disproportionately costly. In cyberspace however, architecture is a function of code; it is the code that defines the conditions of access and the limits of functionality, and code is not static. Programmers can change the code in an instant. The government can also regulate the code by demanding that programmers change it; we cannot say the same about rebuilding bridges. This implies two things for any possible regulatory scheme: first, changes in the code can change the net effect of any policy, and second, it is much easier in cyberspace than it is in real space to direct policies to the architecture.

If we apply Lessig's framework to online speech, we conclude that in the digital ecosystem, the conditions of access to speech and to speakers are both easy and quick to change. This is because it is code that provides the architecture through which speech is regulated online. Programmers can change the access parameters—the code—at their initiative, or because legal regulation requires them to do so. Any *legal* intervention

---

43. *The Law of the Horse*, *supra* note 42, at 507; *see also* CODE AND OTHER LAWS, *supra* note 42, at 30–42.
44. *The Law of the Horse*, *supra* note 42, at 507.

can be directed to this architecture, to the content (speech), or to both.

Let us look at the fake news example in order to contextualize the prior discussion: a problem like fake news is very challenging for regulators. For starters, many would argue that fake news has always existed; it is a very old problem and there is nothing novel about it that requires the attention of the regulators.[45] Further, any attempt to regulate fake news can be perceived as worse than the problem itself.[46] Let us imagine a regulation making fake news illegal and requiring platforms to remove it; in fact, the Malaysian government recently passed such a law.[47] In the U.S., it is extremely unlikely that it could pass First Amendment scrutiny—and for good reason. From a policy perspective, we do not want the government deciding what counts as fake, nor do we want to punish people for being wrong on the Internet, as both can have a chilling effect on speech in general. What regulators *can* do, however, is address the way fake news is spread and amplified through the platforms' architecture. This type of regulation is addressed to the content navigation algorithms and it is not subject to the same doctrinal and normative limitations. It does not target the existence of fake news, but rather the ways in which an old problem has become more salient due to the unique features of the digital ecosystem that control speech online.[48]

---

45. *See* Eugene Volokh, *Fake News and the Law, from 1798 to Now*, WASH. POST: VOLOKH CONSPIRACY (Dec. 9, 2016), https://www.washingtonpost.com/news/volokh-conspiracy/wp/2016/12/09/fake-news-and-the-law-from-1798-to-now/?utm_term=.660884b067ba.

46. Anthony L. Fisher, *Fake News Is Bad. Attempts to Ban It Are Worse*, VOX (July 5, 2017*)*, https://www.vox.com/the-big-idea/2017/7/5/15906382/fake-news-free-speech-facebook-google.

47. Hannah Beech, *As Malaysia Moves to Ban 'Fake News,' Worries About Who Decides the Truth*, N.Y. TIMES, April 3, 2018, at A5; Sandeep Gopalan, *Free Speech Cannot Be Sacrificed to Strike Fake News*, THE HILL (Apr. 6, 2018, 11:00 AM), http://thehill.com/opinion/cybersecurity/381871-free-speech-cannot-be-sacrificed-to-strike-fake-news; Yantoultra Ngui, *Malaysia Passes 'Fake News' Law that Critics Call an Assault on Speech*, WALL ST. J. (April 3, 2018, 8:07 AM), https://www.wsj.com/articles/malaysia-passes-fake-news-law-that-critics-call-an-assault-on-speech-1522757225.

48. Nabiha Syed*, Real Talk About Fake News: Towards a Better Theory for Platform Governance*, 127 YALE L.J. F. 337 (2017), http://www.yalelawjournal.org/forum/real-talk-about-fake-news.

124   *PACE LAW REVIEW*   Vol. 39.1

The fake news example will be used throughout the paper as a thread to demonstrate the asserted arguments. The next section explores the conditions of speech in the digital ecosystem, highlighting why there is a need for regulation under these new conditions.

B. The New Environment

Up until quite recently, the main barrier to speech *entering* the marketplace of ideas was thought of as the (in)ability to publish and access the distribution channels. Content was scarce, and the focus of regulatory schemes was to ensure that it was not suppressed. For instance, the fairness doctrine, which required broadcasters to present both sides of issues of public importance, was valid at a time when broadcast spectrum was scarce. The Court saw the doctrine as promoting First Amendment values[49] such as an unfettered marketplace of ideas.

Today, we live in a very different information environment. Technology has made it extremely cheap, quick, and easy, for just about anyone to create content and make it available online. Platforms, acting as online intermediaries, provide speakers with access to a large audience, having little incentive themselves to monitor the content that is distributed through them.[50] Information has become abundant; in fact, we are drowning in it, and what is now becoming scarce is that which information consumes, i.e. the attention of the listeners.[51] The term *attention economy*[52] has almost made its way into pop

---

49. Red Lion Broad. Co. v. FCC, 395 U.S. 367, 390 (1969).

50. *See infra* Part II(A) (discussing Section 230 of the Communications Decency Act).

51. *See* Herbert A. Simon, *Designing Organizations for an Information-Rich World*, *in* COMPUTERS, COMMUNICATIONS, AND THE PUBLIC INTEREST 37, 40–41 (Martin Greenberger ed., Johns Hopkins Press 1971) (stating "in an information-rich world, the wealth of information means a dearth of something else: a scarcity of whatever it is that information consumes. What information consumes is rather obvious: it consumes the attention of its recipients. Hence a wealth of information creates a poverty of attention and a need to allocate that attention efficiently among the overabundance of information sources that might consume it").

52. *See* TIM WU, THE ATTENTION MERCHANTS: THE EPIC SCRAMBLE TO GET INSIDE OUR HEADS 6 (Knopf 2016).

culture, with more and more news articles referring to the concept.[53]

In this information-rich environment, platforms have taken on the role of the curators, using their algorithms to distribute and deliver content—speech—to their audience. We the listeners need someone—the algorithms—to curate for us and tell us where to look; it is simply impossible to go through everything that is out there on our own. The platforms on the other hand, want to keep our attention because they want to keep us on the platform. The more time users spend on the platform, the more appealing that platform becomes to advertisers, who are then willing to spend more money to capture the attention of these users. For the most part, the revenue of the platforms comes from advertising, so they need to establish a reputation of keeping their audience happy and engaged; this makes users return to the platform and advertisers willing to pay money for those users.

The tricky part is that the platforms' goal is not necessarily welfare enhancing. They just need to keep users *on* the platform. In fact, sometimes lower quality or easier-to-consume-content can serve this goal better.[54] The exact criteria each platform chooses are mostly opaque, but what is becoming clear is that these private companies have developed intricate

---

53. Charles M. Blow, Opinion, *Trump's Attention Economy*, N.Y. TIMES (Jan. 4, 2018), https://www.nytimes.com/2018/01/04/opinion/trump-attention-economy.html; Mathew Ingram, *The Attention Economy and the Implosion of Traditional M*edia, FORTUNE (Aug. 12, 2015), http://fortune.com/2015/08/12/attention-economy/; Andrew Keen, *The 'Attention Economy' Created by Silicon Valley Is Bankrupting Us*, TECHCRUNCH (July 30, 2017), https://techcrunch.com/2017/07/30/the-attention-economy-created-by-silicon-valley-is-bankrupting-us/; Robert Safian, *How to Stand Out in the Attention Economy*, FAST COMPANY (Oct. 3, 2017), https://www.fastcompany.com/40473876/how-to-stand-out-in-the-attention-economy.

54. See PARISER, *supra* note 30, at 68 (describing the theory of "least objectionable programming" as it originates from researching TV viewers' behavior in the 1970s, where it was noticed that with the increasing number of available channels, people quit channel surfing far more quickly than one might suspect. During most of those thirty-six hours a week (that Americans watch TV), the theory suggests, they are not looking for a program in particular, but rather they are just looking to be "unobjectionably entertained."); *see also* SUNSTEIN, *supra* note 30, at 65 (describing information cocoons and group polarization).

126 *PACE LAW REVIEW* Vol. 39.1

systems for moderating and curating speech.[55] As highlighted in the introduction, there is a conceptual distinction between speech curation for hosting purposes versus speech curation for content navigation purposes. The former involves deciding whether content is allowed to exist on the platform, while the latter involves deciding which items to draw the users' attention to.

When it comes to the latter, one could argue that majority tastes and popularity of content have always played a role in content selection and placement.[56] Traditional media has always used these factors in decision-making, especially when choosing headlines, but audience feedback was never as accurate and instantaneous. Big Data technologies now allow for precise tracking and analysis at the moment of content consumption. They also allow for instantaneous adjustment of the content selection based on the feedback. Traditional *editors* never had that kind of power.[57] Algorithmic navigation represents a new kind of knowledge logic, to be contrasted with the *editorial* logic.[58] That it not to say that one is better than the other in

---

55. Klonick, *supra* note 22, at 1660.

56. *See* Robert C. Post, *Data Privacy and Dignitary Privacy: Google Spain, the Right to Be Forgotten, and the Construction of the Public Sphere*, 67 DUKE L.J. 981, 1021, 1023 (2018) (noting that, with the rise of the contemporary American newspaper, "[n]ewspapers began to compete in the 'commodity' of 'news, i.e. information respecting recent events in which the public takes an interest, or in which an interest can be excited'" and "expanded their circulation by reshaping the commodity of news to meet the interests of the masses. Newspapermen justified their expansion by claiming to supply 'what the public wanted—witness their growing sales'") (citations omitted).

57. A/B testing and its application to news headlines is an illustration of how extreme this editing can get. *See A/B Testing*, OPTIMIZELY, https://www.optimizely.com/ab-testing (illustrating how extreme this editing can get through A/B testing and its application to news headlines); *see also* C.W. ANDERSON ET AL., THE NEWS MEDIA: WHAT EVERYONE NEEDS TO KNOW 65 (Oxford Univ. Press 2016) (stating "[b]oth new and old news media also are using digital technology to closely monitor the size and news habits of their audiences, including audiences for individual stories, images, and features on their websites. Some news organizations are using these audience metrics to evaluate the productivity of their journalists and the popularity of their stories, even basing compensation on that data. Some also are using digital traffic data to decide what news to cover, rather than relying only on journalists' news judgment").

58. *See* Gillespie, *supra* note 26, at 192 (noting that "the editorial logic depends on the subjective choices of experts, who are themselves made and authorized through institutional processes of training and certification, or

deciding what information to draw the public's attention to; the point is that they are different.

With significant traffic on news sites coming from social media platforms, editorial decisions of traditional news outlets may now include calculations addressing the ways in which specific platforms filter content. Reporters feel additional pressure to write *click-bait* articles that "pander to readers' worst impulses," as the stories that actually gain traction online are the "[t]oo-good-to-check" stories rather than the ones that are comprehensively reported.[59] In fact, studies show that 59% of the links shared on Twitter have not been read at all by the people who share them.[60] The majority of users simply sees and shares headlines without ever looking at the content of the articles below them. These headlines then get amplified, because in the eyes of content navigation algorithms, the number of shares an article gets is more indicative of what is important and popular than the percentage of users that actually read the article.[61]

In the context of this new algorithmic knowledge logic, traditional news outlets face an uncertain future. They still have to bear the costs of producing content that adheres to journalistic standards, but as their audience has moved to the social web, they have lost control of the distribution channels and the advertising revenues that follow these channels. The platforms are now the biggest distributors of news and, at the same time, they control the public discourse. If we had to draw an (imperfect) analogy, the platforms are both the newsstand

---

validated by the public through the mechanisms of the market. The algorithmic logic, by contrast, depends on the proceduralized choices of a machine, designed by human operators to automate some proxy of human judgment or unearth patterns across collected social traces.")

59. Timothy B. Lee, *Mark Zuckerberg Is in Denial About How Facebook Is Harming Our Politics*, VOX (Nov. 10, 2016, 10:25 PM), http://www.vox.com /new-money/2016/11/6/13509854/facebook-politics-news-bad.

60. MAKSYM GABIELKOV ET AL., SOCIAL CLICKS: WHAT AND WHO GETS READ ON TWITTER? (2016), https://hal.inria.fr/hal-01281190/document; *see also* Caitlin Dewey, *6 in 10 of You Will Share This Link Without Reading It, a New, Depressing Study Says*, WASH. POST (June 16, 2016), https://www .washingtonpost.com/news/the-intersect/wp/2016/06/16/six-in-10-of-you-will-share-this-link-without-reading-it-according-to-a-new-and-depressing-study/?utm_term=.0c9e20635aeb.

61. *Id.*

128   *PACE LAW REVIEW*   Vol. 39.1

and the town square, and the content available on the newsstand is constantly updated based on the town square discussion. Put differently, what used to be just "ephemeral words" in a traditional town square, is now "indexed data" accessible to a remarkable degree of granularity.[62] Additionally, every person sees a different newsstand because content navigation algorithms personalize our news feeds. So the town square is looking less like an *agora* and more like a Tower of Babel.

To put these concerns into perspective, in 2017 two-thirds (67%) of U.S. adults were found to get news on social media (the leader of which is Facebook).[63] The increase in news consumption on social media is especially notable for different demographics than previously: in 2017, 55% of Americans ages fifty or older reported getting news on social media sites versus 45% who reported so in 2016.[64] As audiences move to the mobile and social web, news organizations follow. For instance, CNN, outside its core digital outlets CNN Desktop, CNN Go, CNN mobile web, and CNN apps, uses five different video platforms, twelve different social and messaging platforms, and eleven emerging and off-platforms.[65]

C.   The Dangers

Let us now revisit the fake news example; the phenomenon has attracted global concern. A first-of-its-kind new study by MIT scientists based on Twitter data found that falsehoods "diffused significantly farther, faster, deeper, and more broadly than the truth in all categories of information, and the effects were more pronounced for false political news."[66] A group of

---

62. Syed, *supra* note 48, at 345.

63. *See* Shearer & Gottfried, *supra* note 19, at 2.

64. *Id.*

65. BELL & OWEN, *supra* note 14, at 27 fig.2.

66. Soroush Vosoughi et al., *The Spread of True and False News Online*, 359 SCI. 1146 (2018), http://science.sciencemag.org/content/359/6380/1146 (finding that false news was more novel than true news, which suggests that people were more likely to share novel information); *see also* Robinson Meyer, *The Grim Conclusions of the Largest-Ever Study of Fake News*, ATLANTIC (Mar. 8, 2018), https://www.theatlantic.com/technology/archive/2018/03/ largest-study-ever-fake-news-mit-twitter/555104/.

renowned social scientists are calling for interdisciplinary research to "reduce the spread of fake news and to address the underlying pathologies it has revealed."[67] The reason we are collectively alarmed by the phenomenon is not because it is novel, but because long-standing safeguards against misinformation have eroded.[68] What was previously sitting in the fringes has now inundated public discourse.[69] The agony is not over the *existence* of fake news, but rather over the fact that it can now "spread[] so quickly and persuade[] so effectively."[70] These pathologies of the digital information system are all a function of its architecture. At the center of this architecture are the filtering algorithms that provide content navigation.

In a non-platform world, fake news would get quashed at some point. This is not because anyone would censor it; every citizen of a democracy has an equal right to participate in the public discourse, regardless of whether what he or she has to say is untrue or wrong. Fake news would be quashed in the sense of being irrelevant because in a *functioning* marketplace of ideas, dialogue would eventually lead to a prevailing truth. From a Meiklejohnian perspective, dialogue would ensure that citizens vote in a fully informed and intelligent way, at least in theory.[71] However, in today's platform-dominated world, *dialogue* often takes the form of a continuous reinforcement of existing beliefs within polarized silos of personalization. Now that is a big claim, but the explanation which follows will unpack it.[72]

Filter bubbles and echo chambers were briefly mentioned in the introduction, but it is worth revisiting the concepts in light of all this background.[73] In a world of abundant content and

---

67. Lazer et al., *supra* note 12, at 4.

68. *Id.*

69. *See generally* Syed, *supra* note 48.

70. *Id.* at 337.

71. *See* Robert Post*, Reconciling Theory and Doctrine in First Amendment Jurisprudence*, 88 CALIF. L. REV. 2353, 2367 (2000) (explaining the Meiklejohnian approach as one that understands the First Amendment "to protect the communicative processes necessary to disseminate the information and ideas required for citizens to vote in a fully informed and intelligent way" and views democracy as a process of "the voting of wise decisions." (footnote omitted)).

72. *See also infra* notes 149–178 and accompanying text.

73. *See* Sofia Grafanaki, *Autonomy Challenges in the Age of Big Data*, 27 FORDHAM INTELL. PROP. MEDIA & ENT. L.J. 803, 825 (2017) (discussing

attention scarcity, platforms have to figure out how to best direct users to content. Their economic incentives dictate a need to keep users *on the platform* as long as possible. To accomplish this, they try to serve their users with content that is relevant to them and market this effort as serving the interests of their users. Relevance is measured through complex and opaque calculations involving data on the individual users and users *like them*. Clicks and shares indicate engagement and take priority over the quality of information.[74]

Research has repeatedly illustrated that when confronted with diverse information choices, people tend to choose what feels comfortable and confirms their existing opinion and biases.[75] Individuals are also more likely to believe narratives they have heard before and information that comes from familiar sources, such as friends sharing a link.[76] When consuming information, they also settle for the least objectionable option[77] instead of seeking out the best choice, as the effort alone leads to cognitive overload.[78] Putting everything together, we end up in an environment where users are only directed to content that is familiar to them and in agreement with their existing viewpoints—a filter bubble.[79] On top of that,

---

algorithmic self-reinforcing loops).

74. *See* Matt McGee, *EdgeRank Is Dead: Facebook's News Feed Algorithm Now Has Close to 100K Weight Factors*, MARKETING LAND (Aug. 16, 2013, 9:00 AM), https://marketingland.com/edgerank-is-dead-facebooks-news-feed-algorithm-now-has-close-to-100k-weight-factors-55908; *see also Facebook NewsFeed Algorithm History*, WALLAROO MEDIA https://wallaroomedia.com/facebook-newsfeed-algorithm-change-history/ (last updated Nov. 20, 2018).

75. *See* SUNSTEIN, *supra* note 30, at ch. 4; *see also* Farhad Manjoo, *Our Grip on the Truth*, N.Y. TIMES (Nov. 2, 2016), https://www.nytimes.com/2016/11/03/technology/how-the-internet-is-loosening-our-grip-on-the-truth.html?_r=0.

76. *See* Syed, *supra* note 48, at 350 n.45 (referencing the "illusory truth effect" where "familiarity increases the ease with which statements are processed (*i.e.*, processing fluency), which in turn is used heuristically to infer accuracy").

77. PARISER, *supra* note 30.

78. *See* Margarita Tartakovsky, *Overcoming Information Overload*, PSYCHCENTRAL, https://psychcentral.com/blog/overcoming-information-overload (last updated Jul. 8, 2018) (describing cognitive overload as a phenomenon where our brains get overwhelmed by too much information and too many choices, and effectively freeze, leading to indecisiveness, bad decisions, and stress).

79. *See generally* PARISER, *supra* note 30 (entering the term *filter bubble*

they end up in like-minded communities that transcend previous geographical barriers.[80] When like-minded groups engage in deliberation, they do so in the confines of their *eco-chambers*. The result is a strengthening of the original position and a move towards a more extreme point,[81] amplified by the fact that platforms are designed for fast and frictionless sharing.[82] This kind of deliberation is very different than the one that promotes a democratic culture. There is no exchange of ideas in the marketplace that eventually leads to the truth, or any dialogue in a Meiklejohnian town meeting where decision-making is collective process.[83] On the contrary, mutual understanding between groups becomes harder, leading to group polarization.[84]

This is why the Tower of Babel metaphor was previously used. Facebook's new mission statement stresses an effort to "give people the power to build community. . . strengthen our social fabric and bring the world closer together."[85] So far however, these communities that aspire to be the new *town squares* seem to be doing the opposite; they resemble a world where everyone speaks a different language and they provide fertile ground for falsehoods to spread.

The version of "addictive and toxic misinformation" the world just experienced with the recent fake news epidemic may be just the beginning.[86] Technological tools that can be used to manipulate perception and falsify reality are evolving at a fast pace. The so-called *deepfakes*, i.e. video content that has been manipulated using artificial intelligence, may make the current era of fake news seem antiquated.[87] Skeptics warn of a future

---

to the modern lexicon).

80. Syed, *supra* note 48, at 347.

81. *See* SUNSTEIN, *supra* note 30, at ch. 3 (discussing polarization).

82. Syed, *supra* note 48, at 350.

83. *See* Post, *supra* note 71, at 2367.

84. *See* Del Vicario et al., *supra* note 31.

85. Josh Constine, *Facebook Changes Mission Statement to 'Bring the World Closer Together,'* TECHCRUNCH (June 22, 2017), https://techcrunch.com /2017/06/22/bring-the-world-closer-together/ (internal quotations omitted).

86. Charlie Warzel, *He Predicted the 2016 Fake News Crisis. Now He's Worried About an Information Apocalypse*, BUZZFEED NEWS (Feb. 11, 2018, 8:45 PM), https://www.buzzfeed.com/charliewarzel/the-terrifying-future-of-fake-news?utm_term=.ubzAJl7dY#.plGNq5EMk.

87. Franklin Foer, *The Era of Fake Video Begins*, ATLANTIC (May 15, 2018), https://www.theatlantic.com/magazine/archive/2018/05/realitys-

where our eyes routinely deceive us and anyone could make it appear as if anything has happened, regardless of whether it did.[88] When technology gets there, its existence alone can start impugning content that is real. If anything can be faked, then we reach a point where believing falsehoods and not believing truths are one and the same.[89] Imagine for example being able to make it appear as if a world leader made a statement, which they did not in fact make. When only few individuals have the time or skill to sort truth from fabrication, people give up and stop attempting to figure out what is true and what is fake. The result is a type of *reality apathy* that usually only appears in parts of the world where information is poor and assumed to be incorrect.[90] Hopefully this scenario is just a dystopian nightmare, but one thing we can learn from the recent fake news crisis is that the future comes faster than ever before.

What is true today is the following: our current information environment, as produced by social media content navigating algorithms, has changed the conditions of speech and is not aligned with the prerequisites for citizens in a democratic society to perform their civic duties. Today's information environment undermines the notion of individual autonomy in the selection and consumption of content[91] and threatens the viability of a functioning marketplace of ideas.

First Amendment jurisprudence has been largely shaped by a different information era, namely a time where the audience had plenty of time to hear the speakers and the main threat to the marketplace of ideas was from government censorship of speakers.[92] Today's information environment has very different dynamics than the environment the First Amendment can

---

end/556877/?utm_source=atlfb.

88. *Id.*

89.   Aviv Ovadya, *What's Worse Than Fake News? The Distortion of Reality Itself*, WASH. POST (Feb. 22, 2018), https://www.washingtonpost.com/news/theworldpost/wp/2018/02/22/digital-reality/?utm_term=.b0eaea1c7f81.

90.   Warzel, *supra* note 86.

91.   *See* Grafanaki, *supra* note 73; *see also* Sofia Grafanaki, *Drowning in Big Data: Abundance of Choice, Scarcity of Attention and the Personalization Trap, A Case for Regulation*, 24 RICH. J.L. & TECH. 1, 26–36 (2017).

92.   Tim Wu, *Is the First Amendment Obsolete?*, KNIGHT FIRST AMENDMENT INST. 19 (Sept. 2017), https://knightcolumbia.org/content/tim-wu-first-amendment-obsolete.

protect, yet the doctrinal focus on protecting the speaker remains, which makes it hard to find effective solutions to problems like trolling and fake news.[93] We may all agree that fake news is undesirable, but we do not want the government determining what counts as fake. That is exactly what the First Amendment is supposed to safeguard us from; any regulation trying to address such an issue would be both a content-based and viewpoint-based regulation subject to the strict scrutiny standard, which is very hard to satisfy.

Paradoxically, however, the same doctrine that is meant to safeguard a functioning marketplace of ideas is almost making it harder to do that in today's environment. This is because, as several scholars have recently pointed out, today's tactics for suppressing speech are very different than the ones the First Amendment was envisioned to address.[94] *More speech* and *cheap speech* are the new ways of speech control, used to flood out and drown other speakers, thus undermining them indirectly.[95] Troll armies and bots are often used in this process as well.[96]

Scholars go as far as to question whether the First Amendment is, in fact, obsolete in such an environment[97] and argue that the First Amendment can itself be the barrier to protecting its underlying values.[98] This is because today's speakers face threats, not just from the government, but also from the content curation systems they are filtered through. Yet

---

93. *See generally* Alice Marwick & Rebecca Lewis, *Media Manipulation and Disinformation Online*, DATA & SOC'Y (May 15, 2017), https://datasociety.net/pubs/oh/DataAndSociety_MediaManipulationAndDisin formationOnline.pdf (describing the dynamics of media manipulation and disinformation online).

94. ZEYNEP TUFEKCI, TWITTER AND TEAR GAS: THE POWER AND FRAGILITY OF NETWORKED PROTEST 226 (Yale Univ. Press 2017).

95. *See* Wu, *supra* note 92, at 3; *see generally* Richard L. Hasen, *Cheap Speech and What It Has Done (to American Democracy)*, 16 FIRST AMEND. L. REV. 200 (2017).

96. Individuals have also used more speech as a tool for reputation management purposes, for instance creating multiple content entries about themselves that link to each other in order to *push down* Google Search results that they are trying to *erase*. *See generally* REPUTATION.COM, https://reputation.com (last visited Jan. 29, 2019).

97. *See generally* Wu, *supra* note 92.

98. Jack M. Balkin, *Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation*, 51 U.C. DAVIS L. REV. 1149, 1153–54 (2018).

the platforms designing these algorithmic architectures are not subject to the First Amendment as private entities, but rather use the First Amendment as a shield against attempts to regulate their code. One could again say that we have heard this all before,[99] but we have also seen the Court uphold regulations as constitutional when legally created markets had harmful consequences for free speech.[100] In those cases, the autonomy of the speakers was seen as secondary to the commitment to democratic self-government.[101]

It is for these reasons that the introduction of regulation targeting the underlying architecture is desirable. Within this architecture, if regulation is addressed to the content *navigation* algorithms as opposed to content *moderation* policies, it can be aligned with constitutional values. The next part of this Article starts by setting out the normative and doctrinal barriers that make it hard to regulate content moderation algorithms or impose liability for content that survives through them. It then proceeds to show that these barriers do not apply in the context of content navigation.

---

99. *See* Cass R. Sunstein, *Free Speech Now*, 59 U. CHI. L. REV. 255, 276–77 (1992) (stating "[t]he idea that threats to speech come from government is correct, but as conventionally understood, it is far too simple. Sometimes threats come from what seems to be the private sphere, but those threats are fundamentally a product of legal entitlements that enable some private actors but not others to speak and to be heard. When this is so, these legal entitlements pose a large risk to a system of free expression, one not readily visible to current law.").

100. *See* Red Lion Broad. Co. v. FCC, 395 U.S. 367, 390 (1969) (stating "[i]t is the right of the viewers and listeners, not the right of the broadcasters, which is paramount. It is the purpose of the First Amendment to preserve an uninhibited marketplace of ideas in which truth will ultimately prevail, rather than to countenance monopolization of that market, whether it be by the Government itself or a private licensee. . . . It is the right of the public to receive suitable access to social, political, esthetic, moral, and other ideas and experiences which is crucial here.") (citations omitted).

101. *See generally* Sunstein, *supra* note 99, at 276 (stating "[t]his vision of the First Amendment does not stress the autonomy of broadcasters with current ownership rights. Instead it emphasizes the need to promote democratic self-government by ensuring that people are presented with a broad diversity of views about public issues.").

2018   *REGULATING THE FILTERS*   135

II. Doctrinal Analysis

There is an ongoing effort in the academic literature to find an analogy that places platforms in a category with existing rules and boundaries, so that we can approach them appropriately from a regulatory perspective.[102] Candidates for this analogy include editors, newspapers, broadcasters, cable companies, public utilities, monopolies, state actors, and public forums. However, analogies are always imperfect, and there is no universal agreement on which of the analogies best suits platforms. Some scholars point out the ways in which the platforms' role is explicitly editorial, thus deserving the highest First Amendment protection,[103] while others point out how the editorial analogy fails and these entities are best thought of as cable providers or public utilities.[104]

Platforms perform multiple functions and take on different roles to do so; if we place platforms in a single category—one role, one function—we oversimplify their complex nature. Only by analyzing these roles and functions separately and independently from one another, we can come up with a regulatory framework that can make sense, both doctrinally and normatively. Legal frameworks that apply to some of these functions do not apply to other functions. In some cases, the platforms' algorithms perform (something like) an editorial function resembling a newspaper editor's judgment about what to publish and what not to publish.[105] Those types of algorithms

---

102. *See* Whitney, *supra* note 10, at 30.

103. *See generally* Eric Goldman, *Of Course the First Amendment Protects Google and Facebook (and It's Not a Close Question)*, KNIGHT FIRST AMENDMENT INST. (Feb. 2018), https://knightcolumbia.org/content/course-first-amendment-protects-google-and-facebook-and-its-not-close-question.

104. *See generally* Danah Boyd, *Facebook Is a Utility; Utilities Get Regulated*, DANAH BOYD: APOPHENIA (May 15, 2010), http://www.zephoria.org /thoughts/archives/2010/05/15/facebook-is-a-utility-utilities-get-regulated.html; K. Sabeel Rahman, *Monopoly Men*, BOS. REV. (Oct. 11, 2017), http://bostonreview.net/science-nature/k-sabeel-rahman-monopoly-men; Cale Guthrie Weissman, *Maybe It's Time to Treat Facebook Like a Public Utility*, FAST COMPANY (May 1, 2017), https://www.fastcompany.com/40414024/ maybe-its-time-to-treat-facebook-like-a-public-utility.

105. Zhang v. Baidu.com, Inc., 10 F. Supp. 3d 433, 438–39 (S.D.N.Y. 2014) (finding the same for search engines); *see also* e-ventures Worldwide, LLC v. Google, Inc., Case No. 2:14–cv–646–FtM–PAM–CM, 2017 WL 2210029, at *4 (M.D. Fla. Feb. 8, 2017).

cannot and should not be regulated. In other cases, their algorithms perform functions that are not editorial; the First Amendment should not preclude us from regulating those to an extent. However, if we try to place platforms in strictly defined categories, we will be solving one issue and creating another.

This is why the distinction between content navigation algorithms and algorithms that implement the platforms' content moderation policies is so important. The algorithms that implement content moderation policies can be viewed as being the closest to the editorial analogy, arguably deserving First Amendment protection.[106] On the other hand, the content navigation algorithms can be viewed as non-speech for First Amendment purposes. Further, even if we were to view content navigating algorithms as some type of protected speech, it is argues that this would be non-political speech and, therefore, less problematic to regulate.

Before diving into the doctrinal analysis however, a clarification is in order: while it may seem counterintuitive, the idea that computer code is a type of speech has received considerable support.[107] But what do we mean when we say that an algorithm is protected speech for First Amendment purposes? Whose speech is that? As Tim Wu has observed: "computers make trillions of invisible decisions each day; the possibility that each decision could be protected speech should give us pause."[108] Machine speech and the rights of robots are now topics that have entered the public debate,[109] but in the context of this paper, what is examined is whether the algorithms embody the expression of their designers in a way that deserves constitutional protection. A book embodies the expression of its

---

106. To be clear, I am not advocating for First Amendment protection in the case of content moderation; I am simply refraining from arguing that the First Amendment does not apply.

107. *See generally* Eugene Volokh & Donald M. Falk, *Google: First Amendment Protection for Search Engine Search Results*, 8 J.L. ECON. & POL'Y 883 (2012); *see also* Goldman*, supra* note 103.

108. Tim Wu, *Free Speech for Computers?*, N.Y. TIMES (June 19, 2012), https://www.nytimes.com/2012/06/20/opinion/free-speech-for-computers.html.

109. Tim Wu, *Machine Speech*, 161 U. PA. L. REV. 1495, 1496 (2013); *see generally* James Grimmelmann, *Speech Engines*, 98 MINN. L. REV. 868 (2014); Toni M. Massaro, Helen Norton & Margot E. Kaminski, *Siri-ously 2.0: What Artificial Intelligence Reveals About the First Amendment*, 101 MINN. L. REV. 2481 (2017).

author and, as such, it is protected speech; can we say the same of an algorithm?

As is the case with most hard questions, the answer is: it depends; sometimes yes and sometimes no. As Jack Balkin points out, the deeper issue is "whether companies will be able to shield themselves from regulation by claiming that their uses of AI agents, robots, and algorithms are First Amendment protected activities."[110] Ultimately, what we are trying to figure out is "which business practices are shielded—and should be shielded—from government regulation."[111]

The next two sections address the two types of algorithms that platforms use to curate content separately—moderation vs navigation—and elaborate on the reasoning for treating them differently.[112] The section that follows makes a further argument limited to personalization algorithms, which is that personalization algorithms may not even deserve the limited protections available for non-political protected speech, as they are better categorized as professional or commercial communications that fall outside the First Amendment protection altogether.[113]

A.   Content Moderation Policies and Section 230 of the Communications Decency Act

Content-moderating practices that monitor for unlawful or other undesirable content are viewed as falling under the realm of Section 230,[114] giving platforms the freedom to moderate content without facing risks of liability. There is a big debate around Section 230 of the Communications Decency Act ("CDA"), its effect on online speech, and its desirability; however, in the context of this paper, it suffices to explain its origins and perceived purpose. The main idea is that online intermediaries are not considered publishers or speakers of information provided by "another information content provider."[115] In other

---

110.   Balkin, *supra* note 98, at 1159.
111.   *Id.* at 1160.
112.   *See infra* Parts II(A), (B).
113.   *See infra* Part II(C).
114.   47 U.S.C. § 230 (2012).
115.   *Id.*

words, they are not liable for user-generated content. The reasoning is that if the intermediaries could be found liable for defamatory content posted by a user, they would have to take an active role in policing their users' content and would naturally err on the side of censorship. That was seen as a danger to freedom of speech online, and Section 230 prevented that from happening. Because of the ways *publisher* and *intermediary* are defined, Section 230 goes even further, so as to protect the good Samaritans that actually try to take some precautions against defamatory content.

Regardless of whether this is good or bad law, the end result is that platforms, as online intermediaries, while not obligated to monitor and police content provided by others using the platform, are at the same time free to do so under their own rules. Recent accounts by commentators illustrate just how closely these moderation systems resemble a "legal system;" they curate user content "with an eye to American free speech norms, [and] corporate responsibility."[116] Yet they operate outside of the boundaries of the law *per se*. At their core, platforms are driven by their economic interests, which dictate a "necessity [to] creat[e] an environment that reflects the expectations of their users."[117]

One could argue that the fact that Section 230 grants immunity from liability to the good Samaritans is not tantamount to a First Amendment right of free expression. Just because platforms are not liable as publishers for the content they moderate, it does not necessarily follow that their moderation practices deserve constitutional protection. However, out of all the different functions that machine-learning algorithms perform, these content moderation algorithms are the closest to an editorial function. Recall that content moderation algorithms determine whether content is allowed to exist on the platform, as opposed to where it is placed and how it is promoted, which is the subject of content navigation. The decision whether to allow or disallow content is based on the *substance* and the *message* of the content, not on clicks or other metrics. Humans are continuously involved in this process, and

---

116. Klonick, *supra* note 22, at 1599, 1664.
117. *Id.* at 1599, 1602, 1669.

they sometimes override the algorithms.[118]   Platforms are choosing the type of content that reflects the environment they want to create; interfering with this process can be seen as a close analogy to compelled speech.[119]

In fact, two lower courts' decisions have come to this conclusion in the context of search engines where Baidu and Google had delisted websites from their search results. The idea is that, just as the First Amendment protects newspaper editors who cannot be compelled to publish a particular content item, the same concept applies to search engines, which cannot be compelled to include certain links.[120]   The decisions of search engines to list or delist content as part of their index is the closest equivalent to content moderation policies of social media platforms; they both deal with the question of whether content is allowed to exist—on the index or the platform. Both in *Zhang v. Baidu.com, Inc.*[121] and in *e-ventures Worldwide, LLC v. Google, Inc.*,[122] the Courts considered *Miami Herald Publishing Co. v. Tornillo* to be the governing precedent; a case where the Supreme Court held that requiring newspapers to provide a right to reply to political candidates constituted compelled speech and was unconstitutional.[123]   In *e-ventures*, the Court found that "[a] search engine is akin to a publisher, whose judgments about what to publish and what not to publish are absolutely protected by the First Amendment."[124]   Google has also won lawsuits on free expression grounds against claims that challenged its choice of which advertisements to display.[125] In *Langdon v. Google, Inc.*, the District Court rejected a claim that Google had a duty to carry specific advertisements, and again found that Google's advertising decisions were similar to those

---

118. *See generally id.*

119. *See* Miami Herald Publ'g Co. v. Tornillo*,* 418 U.S. 241, 252 (1974).

120. *See generally* Volokh & Falk, *supra* note 107.

121. 10 F. Supp. 3d 433, 436 (S.D.N.Y. 2014).

122. Case No. 2:14–cv–646–FtM–PAM–CM, 2017 WL 2210029, at *3–4 (M.D. Fla. Feb. 8, 2017).

123. *Tornillo*, 418 U.S. at 258.

124. *e-ventures*, 2017 WL 2210029, at *4 (citing *Tornillo*, 418 U.S. at 258).

125. *See generally* Langdon v. Google, Inc., 474 F. Supp. 2d 622 (D. Del. 2007) (displaying the ease with which lower courts have applied the editorial analogy). I am noting that advertising decisions are of course different than decisions about listing/delisting content.

of a newspaper and, as such, could not be regulated.[126]

Scholars have questioned the logic of these decisions. Frank Pasquale notably views the courts' reliance on *Tornillo* as misguided and argues that there is a difference between trying to advertise in one out of hundreds of newspapers versus in *the one* dominant search engine.[127] For Pasquale, to the extent that Google is a media entity, it is closer to the entities subject to the fairness doctrine in *Red Lion*;[128] he sees a need to ensure "platform neutrality" in order to prevent intermediaries from distorting the public sphere or private commerce by virtue of their size and dominant position.[129]

Social media platforms' content moderation decisions may or may not be distinguishable from search engines' listing, delisting, and choice-of-advertisement decisions. The functions are not identical, but, for now, these are the closest judicial takes on hosting decisions. Lower courts have been quick to apply the editorial analogy, finding that search engines' decisions on what content to list or delist, or which advertisements to carry, are similar to a publisher's decisions about what to publish and, therefore, fully protected by the First Amendment.

While there are arguments that support the opposite conclusion[130] and the platforms repeatedly make a point to disassociate themselves from the idea that they are acting as editors or newspapers, ultimately, it may be best from a policy perspective not to interfere with the way content moderation systems work. If, for instance, we viewed them as state actors whose users had First Amendment rights, as adopting a different analogy would suggest, the result would be that platforms would be unable to police spam. They could be quickly overwhelmed by other undesirable content and to avoid "mass exodus" of their users they would probably redesign their algorithms in a way that is explicitly more editorial so that they

---

126. *Id.* at 630 (citing *Tornillo*, 418 U.S. at 256).

127. Frank A. Pasquale, *Platform Neutrality: Enhancing Freedom of Expression in Spheres of Private Power*, 17 THEORETICAL INQUIRIES L. 487, 502–03 (2016).

128. *Id.* at 503; *see generally* Red Lion Broad. Co. v. FCC, 395 U.S. 367 (1969).

129. Pasquale, *supra* note 127, at 489.

130. *See* Whitney, *supra* note 10, at 8–13 (discussing the internal weaknesses of the editorial analogy).

can claim First Amendment protection.[131] However, doing so would take them outside of the realm of Section 230, and, in order to avoid liability for third party content, they would err on the side of caution and censorship. Alternatively, if the government were to dictate specific rules for content moderation (for example, fake news is illegal and should be banned), then we have the government in the position of deciding what is *fake*, resulting in indirect censorship because of its chilling effect.

So, regardless of which analogy applies best for First Amendment purposes, it is not suggested that regulators interfere with the content moderation systems. Even if we could agree on the analogy in this context today, it may not apply tomorrow. Going back to Lessig's point on the plasticity of the architecture of cyberspace,[132] platforms can very easily change the code and make today's analogy inapplicable tomorrow.

The next section discusses content navigation algorithms, and argues that they should be subject to different treatment and analysis.

B.   Content Navigation Algorithms in General

Content navigation algorithms can be described as machine-learning algorithms that continuously adjust themselves based on new data inputs. These types of algorithms should not qualify as speech for First Amendment purposes and, even if they do, they are lower value speech as opposed to political speech. As such, any government regulation that addresses them should be assessed under a lower standard of scrutiny.

In terms of judicial precedent, the notion that the First Amendment protects content navigation code has again been made in the context of Google's search algorithm. The search algorithm is the closest equivalent to platform's content navigation algorithms because the question it addresses is not whether content exists, but rather where it is placed and whether the user's attention is directed to it. In *Search King, Inc. v. Google Technology, Inc.*,[133] the District Court was faced

---

131. Goldman, *supra* note 103 (internal quotation marks omitted).
132. *See supra* Part I discussion and accompanying footnotes.
133. No. CIV-02-1457-M, 2003 WL 21464568, at *1 (W.D. Okla. May 27,

142   *PACE LAW REVIEW*   Vol. 39.1

with the issue of "whether a representation of the relative significance of a web site as it corresponds to a search query is a form of protected speech."[134] Google argued that PageRanks are subjective opinions, and the Court further accepted that search results are analogous to financial ratings. It then relied on the Tenth Circuit's decision in *Moody's Investor's Services*, where Moody's rating of bonds was found to be "a statement of opinion relating to matters of public concern which does not contain a provably false factual connotation" and as such received "full constitutional protection."[135] Based on the analogy, the *Search King* Court concluded that PageRanks are "opinions of the significance of particular web sites as they correspond to a search query," that they "relate to matters of public concern," and because there is "no conceivable way to prove that the relative significance assigned to a given web site is false," Google's PageRanks are entitled to "full constitutional protection."[136]

A notable point in the opinion, however, is the Court's emphasis on the subjective nature of the search *results* as opposed to the objective nature of the *process* performed by the algorithm. The Court used this distinction to place search *results* and not search *algorithms* within the "protected class of speech."[137] In a passage worth quoting in full, the District Judge stated:

> Here, the process, which involves the application of the PageRank algorithm, is objective in nature. In contrast, the result, which is the PageRank - or the numerical representation of relative significance of a particular web site - is fundamentally subjective in nature. This is so because every algorithm employed by every search engine is different, and will produce a

2003).

134. *Id.*

135. Jefferson County Sch. Dist. No. R-1 v. Moody's Inv'r's Servs., 175 F.3d 848, 852 (10th Cir. 1999) (quoting Milkovich v. Lorain Journal Co., 497 U.S. 1, 19–20 (1990)) (internal quotation marks omitted).

136. *Search King*, 2003 WL 21464568, at *3–4 (quoting *Jefferson County*, 175 F.3d at 852) (internal quotation marks omitted).

137. *Id.* at *3.

> different representation of the relative
> significance of a particular web site depending on
> the various factors, and the weight of the factors,
> used to determine whether a web site corresponds
> to a search query. In the case at bar, it is the
> subjective result, the PageRank, which was
> modified, and which forms the basis for Search
> King's tort action.[138]

Regardless of whether this distinction makes conceptual sense, it is important in that it clarifies what *speech* the *Search King* Court felt was deserving of First Amendment protection. It was not the search algorithm, which the Court viewed as an objective process, but rather the search *result*, which took subjective factors into account. Following this logic, *Search King* cannot be read as holding that content navigation algorithms deserve First Amendment protection because the decision was focused solely on the results. That said, the distinction is not as clear-cut as the Court portrays it to be. The factors and the weights of the factors are also, for the most part, products of algorithmic calculations. The more we drill down, the distinction between process and result starts fading, as machine-learning algorithms constantly update their *process* based on *results*. *Search King* is also just a district court decision; so, regardless of the distinction the Court made, we need to take a closer look at the First Amendment issues that arise in the case of social media platforms.

As opposed to Google, who has outright argued that PageRanks are subjective opinions,[139] social media platforms have not yet made an argument that their news feeds or

---

138. *Id.* at *3–4.

139. *See generally id.* Even though Google argues that its PageRanks are subjective opinions, it is important to note that it does not take the same position with respect to other algorithmic functions it performs, such as its auto-complete function for instance. Instead, Google explicitly states that such auto-complete search predictions are not the answers to a search query, but rather the terms are predicted by computer algorithms based on searches from previous users, not by Google itself.); *see also* David Meyer, *Google Loses Autocomplete Defamation Case in Italy*, ZDNET (Apr. 5, 2011, 2:13 PM), http://www.zdnet.com/article/google-loses-autocomplete-defamation-case-in-italy/.

144 *PACE LAW REVIEW* Vol. 39.1

trending news features represent protected speech; they even go out of their way to deny any resemblance to an editorial role. Notably, Greg Marra, the Facebook engineer whose team designs the code that drives Facebook's News Feed, has plainly stated:

> We try to explicitly view ourselves as not editors . . . We don't want to have editorial judgment over the content that's in your feed. You've made your friends, you've connected to the pages that you want to connect to and you're the best decider for the things that you care about.[140]

Presumably, platforms take this stance because being perceived as editors is a double-edged sword. If they are editors and their algorithms are speech, then it follows that they are media companies; but they really *do not* want to be media companies. They want to be seen as *technology* companies who do not produce any original content, but merely distribute it in a neutral way, and whose algorithms simply *reflect* what users want.[141]

Regardless of their self-categorization, there is a conceptual difference between search engine results and social media news feeds. When people use Google, they are searching for something with the expectation of being presented with correct and relevant answers to their inputted keywords and queries. They may use Google as opposed to Bing because of Google's reputation as a better search engine, and the process resembles a type of dialogue where a user actively asks a question and expects Google to come up with the best answer. The user is essentially asking Google for its viewpoint on a particular query. This *dialogue* sometimes continues when the user adjusts search parameters such as time period. In that sense Google's

---

140. Ravi Somaiya, *How Facebook Is Changing the Way Its Users Consume Journalism*, N.Y. TIMES (Oct. 26, 2014), https://www.nytimes.com /2014/10/27/business/media/how-facebook-is-changing-the-way-its-users-consume-journalism.html.

141. *See* Philip M. Napoli, Sanford Sch. of Pub. Policy, & Robyn Caplan, Data & Soc'y Research Inst., When Media Companies Insist They're Not Media Companies and Why It Matters for Communications Policy, (Sept. 30, 2016), https://ssrn.com/abstract=2750148.

algorithm can be seen as expressive and somewhat editorial. Social media companies, on the other hand, are not perceived that way and do not intend to be perceived that way. They explicitly refrain from expressing any viewpoint. Users do not ask Facebook what content Facebook thinks is relevant to them; they use Facebook because that is where their friends are. They individually select which friends to connect with and which pages to follow, and, when users log on to Facebook, they expect to view content updates from the individuals and pages they have themselves selected. The focus is not on Facebook's selection and arrangement of these updates; in fact, this is mostly obscured by the company's statements.[142]

Interestingly, in its recent decision, *Packingham v. North Carolina*, the Supreme Court found social media to be "the modern public square," "the most important places (in a spatial sense) for the exchange of views" and stated that it is a "fundamental principle of the First Amendment . . . that all persons have access" to such a forum.[143] The context of this case was very different, as it concerned a law prohibiting registered sex offenders from social networking platforms, but it is the analogy the Court drew and the stress on the right of access that matters for our purposes. The Court's approach suggests that the role of these platforms in public discourse is entirely different from that of traditional editors and, therefore, should not be treated as such for First Amendment purposes. Requiring all persons to have access to a forum is exactly what the government cannot do when editors are involved.[144]

The editorial speech courts have traditionally protected[145] represents an editor's judgment about which issues are of public importance and a commitment to journalistic ethics. That is, political speech and the First Amendment strongly protects it. In the case of Facebook, even if the actual content selected is political in nature, the navigation algorithm itself is not. The nature of machine-learning algorithms is such that they are

---

142. *See* Somaiya, *supra* note 140.

143. 137 S. Ct. 1730, 1732, 1735, 1737 (2017).

144. *See generally* Miami Herald Publ'g Co. v. Tornillo, 418 U.S. 241 (1974).

145. *See* Columbia Broad. Sys., Inc. v. Democratic Nat'l Comm., 412 U.S. 94, 120 (1973).

constantly changing and adjusting to new inputs in ways that are so complex, that they often go beyond the comprehension of their designers. It would be very odd, to say the least, to grant the same protection we give to political speech to algorithmic processes and outputs that cannot even be explained by their own designers.

There is no particular viewpoint or message embodied in these algorithmic processes; the *expression* they represent is driven entirely by clicks and engagement metrics. They are designed to optimize user engagement. The *content* of the algorithm is constantly changing, and, in the case of personalization algorithms, it is not even directed to more than one person. Functionally, this is more like *selling* than *editing*. Even if the content items that the algorithms filter are political in nature, the algorithms themselves are not political speech. So, if navigation algorithms are protected speech at all, we need to ask what type of protected speech that is.[146]

Doctrinally, speech which does no more than "propose a commercial transaction"[147] or "link . . . a product to a current public debate"[148] is considered commercial speech and deserves a lower degree of constitutional protection. The classic example of commercial speech is advertising, but there have been instances where advertising has been considered public discourse and instances where expression that is not an advertisement has been considered commercial speech.[149] Content navigation algorithms are not a clear example of commercial speech, but they do not need to be in order to justify lower constitutional protection. The commercial speech doctrine is relevant for our purposes because it illustrates *why* and *how* the First Amendment protects non-political speech. Instead of

---

146. *See* Robert Post, *Compelled Commercial Speech*, 117 W. VA. L. REV. 867, 871–72 (2015) (arguing that "First Amendment doctrine protects each distinct kind of speech in a manner appropriate for safeguarding its particular kind of constitutional value. 'Speech as such' does not contain any constitutional value." (footnote omitted)).

147. *See* Va. State Bd. of Pharmacy v. Va. Citizens Consumer Council, 425 U.S. 748, 760 (1976).

148. *See* Cent. Hudson Gas & Elec. Corp. v. Pub. Serv. Comm'n, 447 U.S. 557, 562 n.5 (1980).

149. *See* Robert Post, *The Constitutional Status of Commercial Speech,* 48 UCLA L. REV. 1, 5 (2000) (citing Bolger v. Youngs Drug Prods. Corp., 463 U.S. 60 (1983)); *see generally* N.Y. Times Co. v. Sullivan, 376 U.S. 254 (1964).

focusing on which descriptive category of speech content navigation algorithms fall under, it is best to focus on the constitutional value of the expression in question in order to determine whether they deserve constitutional protection at all and, if so, to what degree.[150]

Scholars like Edwin Baker have argued that commercial speech should not be protected at all.[151] This approach reflects an autonomy theory of free speech, meaning that individual liberty and personal agency require speakers to be free to choose the content of their own message without interference. In the marketplace, however, speech is dictated by profit-maximizing and efficiency standards and, as such, does not represent an exercise of freedom.[152]

From a slightly different perspective, Robert Post views First Amendment doctrine as having not one single structure, but rather as being plural.[153] He explains that different types of speech embody different constitutional values and as such, each type of speech should receive the protection that is appropriate for the values it embodies.[154] Ordinary First Amendment doctrine, meaning doctrine concerning political speech, focuses on the *speaker*. It protects the right of the speaker to participate in public discourse and the equal right of every citizen in a democracy to participate in the formation of public opinion.[155] In contrast, commercial speech is protected for the informational value it provides to the listeners, and commercial speech doctrine is focused on the free flow of information.[156] Robert Post identifies this value as "democratic competence," referring to the "cognitive empowerment of those who participate in public discourse," which is achieved when *listeners* can receive information

---

150. Post, *supra* note 149, at 11.

151. C. Edwin Baker, *The First Amendment and Commercial Speech*, 84 IND. L.J. 981, 997 (2009).

152. *Id.* at 986.

153. Post, *supra* note 146, at 871; Robert Post & Amanda Shanor*, Adam Smith's First Amendment*, 128 HARV. L. REV. F. 165, 181 (2015).

154. Post & Shanor, *supra* note 153, at 181–82.

155. *Id.* at 170.

156. *Id.* (citing First Nat'l Bank v. Bellotti, 435 U.S. 765, 783 (1978).

in order to form their own opinions and make decisions based on them.[157]

In other words, political speech is speech that is itself part of public discourse, and speakers receive the strongest protection when they engage in public discourse. Commercial speech, on the other hand, is not itself part of public discourse, but rather it facilitates the free flow of information so that participants in the public discourse can form their opinions and the content of their (political speech) in an informed and intelligent way.

Against this theoretical background, let us now revisit content navigation algorithms. Even if we accept that content navigation algorithms are speech for First Amendment purposes, they are certainly not speech that is part of the public discourse. Their function is to facilitate the flow of information to listeners. The listeners are the ones that become speakers participating in public discourse; the algorithms are merely information conduits. As such, if they deserve any First Amendment protection against regulation, it is because of the value they provide to the listeners in delivering information. They are different than commercial speech because commercial speech *is itself* the information that has value for the listeners, whereas content navigation algorithms are merely the *delivery mechanism* of the information. However, they are similar in that the constitutional focus is on the listeners and the listeners' interest in receiving information.

As a new category, content navigation algorithms have no established legal test which can be applied to potential restrictions on them. The closest applicable test is the test for assessing the constitutionality of restrictions on commercial speech. However, courts need not be bound by it, and there is no reason to use an already complex and unclear standard for a category it was not designed for. That said, for the purposes of this paper, it is useful as an exercise to draw on the commercial speech doctrine, as it can highlight the types of issues that will arise. Given that the reasons to grant some level of constitutional protection to content navigation algorithms are similar to the reasons why we protect commercial speech

---

157. *Id.* at 170 n.35 (citing ROBERT POST, DEMOCRACY EXPERTISE, AND ACADEMIC FREEDOM: A FIRST AMENDMENT JURISPRUDENCE FOR THE MODERN STATE 27–60 (Yale Univ. Press 2012)).

(meaning, the informational value it has for the listeners), any new test would presumably draw on ideas similar to those embodied in the test for commercial speech.

The leading case that establishes the test for the constitutionality of restrictions on commercial speech is *Central Hudson Gas & Electric Corp. v. Public Service Commission of New York*, which lays out a four-prong test.[158] First, for the commercial expression to be eligible for First Amendment protection, it must concern lawful activity and not be misleading.[159] Second, it must be determined whether the government interest to be served by the restriction on commercial speech is *substantial*.[160] Third, we ask whether the regulation directly advances the governmental interest asserted; and fourth, whether it is not more extensive than is necessary to serve that interest.[161]

Let us now apply the *Central Hudson* test to the content navigation algorithms of social media platforms. If we assume that the algorithms qualify for protection under the first prong,[162] we then have to articulate the government interest that the regulation would assert and determine if it is a *substantial* interest. Enhancing welfare is typically considered a substantial government interest,[163] and the connection between the circulation of information and welfare is precisely why the doctrine of commercial speech exists.[164] Therefore,

---

158. 447 U.S. 557, 566 (1980).

159. *Id.*

160. *Id.*

161. *Id.*

162. There is an argument to be made that these algorithms are actually misleading and, as such, do not even qualify for limited First Amendment protection available to commercial speech. Selecting fake news items as trending or as relevant to a user may be very misleading, but ultimately that depends on what the algorithms are perceived to do. Because this part of the test can go both ways depending on interpretation, I proceed to examine the remaining three prongs as well.

163. *See* Post*, supra* note 146, at 891.

164. *See id.* (quoting Va. State Bd. of Pharmacy v. Va. Citizens Consumer Council, 425 U.S. 748, 760, 765 (1976) (stating "[s]o long as we preserve a predominantly free enterprise economy, the allocation of our resources in large measure will be made through numerous private economic decisions. It is a matter of public interest that those decisions, in the aggregate, be intelligent and well informed. To this end, the free flow of commercial information is indispensable.").

regulating content navigation algorithms in a way that promotes the free flow of information and encourages the existence of a functioning marketplace of ideas is perfectly constitutional. Doing so promotes true autonomous choice in the selection and consumption of content and the ability to make intelligent and informed decisions. As such, it is aligned with the prerequisites for a "free enterprise economy"[165] and with the requirements for citizens in a democratic society to be able to perform their civic duties, and is, therefore, welfare enhancing.

This leaves us with the last two prongs of the *Central Hudson* test: whether the regulation directly advances the governmental interest asserted, and whether it is more extensive than necessary to promote that interest.[166] The assessment of the last two prongs must be done against the precise language of the regulation, and designing it is inevitably a multi-disciplinary task requiring the collaboration of platforms, engineers, academics, and regulators. However, to illustrate the viability and constitutionality of regulation, an example of the types of requirements that could be imposed on the content navigation algorithms is offered. We can imagine, for instance, a requirement that trending topic algorithms exclude bot activity from measures of what is trending,[167] or a requirement that algorithms treat links that are shared, but unopened, differently from links that have actually been opened. We can also imagine a requirement that content that originated from an ad, even if subsequently shared organically, cannot be included in popularity measures of news; or a requirement to exclude from the news feed algorithms information on how users interact with ads.[168] These are merely examples, because any

---

165.   *Id.* at 891 (citing *Va. State Bd.*, 425 U.S. at 765).

166.   Cent. Hudson Gas & Elec. Corp. v. Pub. Serv. Comm'n, 447 U.S. 557 (1980).

167.   This is in fact a step that platforms have purported to initiate on their own: *See* Colin Crowell, *Our Approach to Bots and Misinformation*, TWITTER: BLOG (June 14, 2017), https://blog.twitter.com /official/en_us/topics /company/2017/Our-Approach-Bots-Misinformation.html; Jen Weedon et al., *Information Operations and Facebook*, FACEBOOK (2017), https://fbnews roomus.files.wordpress.com/2017/04/facebook-and-information-operations-v1.pdf.

168.   *See* McGee, *supra* note 74 (noting that while "the News Feed algorithm is completely separate from the algorithm that decides what ads to show, when to show ads, and where to show them[,] . . . how a user interacts

regulatory proposal has to be the product of systematic interdisciplinary research that evaluates not only the issues, but also the potential effectiveness of the interventions. What the preceding analysis simply illustrates is that the First Amendment does not prohibit such regulatory intervention.

C.   The Special Case of Personalization: Protected Speech v. Commercial Communications

The preceding discussion relates to content navigation algorithms used by social media platforms in general; it includes both algorithms that select content that is *trending* and personalization algorithms and argues that regulating them can be consistent with First Amendment doctrine. This Section will further argue that personalization algorithms specifically may not even deserve the limited constitutional protections of the type afforded to non-political speech.

To understand how to draw the boundaries between different types of speech, it is again illustrative to look at the commercial speech doctrine. The category of commercial speech is defined by two boundaries: the first distinguishes it from public discourse, which is what has been discussed above, and the second falls on the other side of the spectrum and distinguishes commercial speech from other types of communications that do not receive any of the constitutional protections afforded to commercial speech.[169]

Doctrinally, core First Amendment protection involves *political speech*, meaning speech that is part of the public discourse. Non-political speech such as commercial speech, which gets limited protection, involves speech that conveys information to those participating in public discourse. Left unprotected are "those forms of commercial communications that do not serve to underwrite a public communicative sphere."[170] Robert Post's work can again shed light to this distinction. He points out that, for sociologists, this public communicative sphere consists of a shared "universe of

---

with Facebook ads *can* influence what shows in the News Feed").
    169.   *See* Post, *supra* note 149, at 15–25.
    170.   *Id.* at 22.

discourse" which exposes strangers presumed to be "independent and rational" to "similar social stimuli."[171] To illustrate what falls outside of this public communicative sphere, Post uses the paradigmatic example of fiduciary relationships, such as the relationship of lawyers with their clients or doctors with their patients, which even though involve the conveying of information, can be regulated without involving the First Amendment commercial speech doctrine. He suggests that there are implicit assumptions within the doctrine that distinguish between personal communications which constitute relationships of dependence and reliance and impersonal communications addressed to independent and rational citizens.[172]

Personalization algorithms may not be the paradigmatic example of fiduciary relationships of dependence and reliance. That said, scholars have suggested that certain types of online service providers take on fiduciary responsibilities in the digital age. Neil Richards and Jack Balkin have both written on the concept of information fiduciaries in the digital age, arguing that online service providers have a unique relationship with end-users that placed users in a position of dependence and vulnerability.[173] These relationships, they claim, are not identical to the traditional kinds of fiduciaries, such as lawyers and doctors, that would require the strictest fiduciary obligations, but the law should recognize new kinds of fiduciary relations and obligations that correspond to the respective user-entity relationship.[174]

---

171. *Id.* (citing JOHN W. BENNETT & MELVIN M. TUMIN, SOCIAL LIFE: STRUCTURE AND FUNCTION: AN INTRODUCTORY GENERAL SOCIOLOGY 140 (Knopf 1948); Carroll D. Clark, *The Concept of the Public*, 13 SW. SOC. SCI. Q. 311, 313 (1933); Robert Post, *Recuperating First Amendment Doctrine*, 47 STAN. L. REV. 1249, 1276 (1995)).

172. Post, *supra* note 149, at 24.

173. *See* NEIL RICHARDS, INTELLECTUAL PRIVACY: RETHINKING CIVIL LIBERTIES IN THE DIGITAL AGE ch. 10 (Oxford Univ. Press 2015); Jack M. Balkin, *Information Fiduciaries and the First Amendment,* 49 U.C. DAVIS L. REV. 1183, 1221–22 (2016).

174. Balkin, *supra* note 173, at 1223 (stating "we should recognize that a changing society generates new kinds of fiduciary relations and fiduciary obligations that the law can and should recognize. The scope of the fiduciary duty, however, is not the same for every entity. It depends on the nature of the relationship, the reasonableness of trust, and the importance of preventing self-dealing by the entity and harm to the end-user, client, or beneficiary."); *see*

To be clear, this Article is not suggesting that the special duties of the information fiduciaries of the digital age extend to the way personalization algorithms work. Both Balkin and Richards are mostly focused on privacy and the handling of personal information when they talk about these new fiduciary obligations. This Article *is* however suggesting, that the nature of the relationship between platforms and individual end-users implies that communications in the context of this relationship fall outside of public discourse altogether. In other words, these communications should not even receive the limited protections available to non-political speech, such as commercial speech.

In *Lowe v. SEC* for instance, the Court found that "personalized communications create special dangers of 'fraud, deception, or overreaching' that 'are not replicated in publications that are advertised and sold in an open market,'" and found that Congress can treat investment advisors as fiduciaries consistent with the First Amendment.[175] Regardless of whether fiduciary responsibilities should be imposed on platforms, what is argued is that, on the personalized level, they perform a function that falls outside the realm of constitutional speech protections. This is especially so when there are reasons to doubt the autonomy of the listeners.[176]

---

also RICHARDS, *supra* note 173, at 168 (stating "[j]ust as we recognized in the past that certain professionals were fiduciaries of our information, so, too, in the Age of Information should we expand our definition of information fiduciaries to include bookstores, search engines, ISPs, email providers, cloud storage services, providers of physical and streamed video, and websites and social networks when they deal in our intellectual data. The duties of confidentiality we place on these fiduciaries need not be ironclad. Sometimes we want to share our views with the world, and intermediaries can help to do that, too. But that should be our choice, not theirs.").

175. *See* Balkin, *supra* note 173, at 1219 (citing Lowe v. SEC, 472 U.S. 181, 210 (1985). Jack Balkin clarifies that the Court did not say that investment advisors *must* be treated as fiduciaries, but simply said that the First Amendment does not preempt Congress from doing so. *Id.*

176. *See* Post, *supra* note 149, at 41 (offering yet another approach for why these types of fiduciary relationships do not deserve First Amendment protection, viewing them as failing the first prong of the *Central Hudson* test (i.e. the misleading requirement), and suggesting that the requirement should be redefined "to focus on the specific conditions that might be understood to render consumers dependent and vulnerable," rather than focusing on the content of speech).

Let us recall the Facebook social contagion experiment, which caused an outrage in 2012.[177] For a week in January 2012, the feeds of about 700,000 Facebook users were manipulated to determine how users' emotional states change depending on the nature of the posts they see.[178] Some users saw content with a preponderance of happy and positive words, while others saw content that was considered sadder than average.[179] At the end of the week, these users were more likely to post content that corresponded to the type they had been exposed to, that is especially positive or negative words.[180] The experiment concluded that "emotional states can be transferred to others via emotional contagion, leading people to experience the same emotions without their awareness," and that "emotional contagion occurs without direct interaction between people (exposure to a friend expressing an emotion is sufficient), and in the complete absence of nonverbal cues."[181] Despite the public outrage, what Facebook allowed to happen on its platform was probably legal. If so, that would be because of a contractual relationship between Facebook and each user established through terms of service and privacy policies. It is *not* because Facebook has some First Amendment right bestowed upon its News Feed.

The fact is that the proliferation of content coupled with the information asymmetries of the attention economy places users of social media platforms in a position of disadvantage and relative dependence on the platforms. There is so much content out there that individuals do not have the time or attention span to actively select what they engage with, given the plethora of available choices. Curators are needed more than ever, and

---

177. *See* Kate Crawford, *The Test We Can—and Should—Run on Facebook*, ATLANTIC (July 2, 2014), https://www.theatlantic.com/technology /archive/2014/07/the-test-we-canand-shouldrun-on-facebook/373819/; *see also* Robinson Meyer, *Everything We Know About Facebook's Secret Mood Manipulation Experiment*, ATLANTIC (June 28, 2014), https://www.theatlantic .com/technology/archive/2014/06/everything-we-know-about-facebooks-secret-mood-manipulation-experiment/373648/.

178. *See* Adam D. I. Kramer et al., *Experimental Evidence of Massive-Scale Emotional Contagion Through Social Networks*, 111 PROC. NAT'L ACAD. SCI. U.S. AM. 8788, 8788 (2014).

179. *Id.*

180. *Id.*

181. *Id.*

unlike search engines users that search for something specific, social media platform users tend to passively consume content that makes it on their news feeds. The platforms claim that they are giving users what they want, so why, one would ask, is there anything problematic here? Haven't people always chosen to read the newspaper or watch the TV channel that corresponds to their political beliefs? It is not suggested that we need paternalistic guidance on what we *should* be reading, nor that the government is a better judge than the platforms of that is. What is suggested, however, is that personalization is not quite the same as a man choosing to read his party's newspaper; it represents a different knowledge logic.[182]

In personalized news feeds, user preferences are implied based on factors that are opaque. At a minimum, they include actual data the platform has on the individual, presumably with the individual's consent.[183] But they also include inferences that platforms make about the users, which may or may not be accurate. These isolated points of information (actual data and inferences) are used to construct what has been called our "data doubles"[184] or "shadow bodies," which emphasize some characteristics (in the form of data points), and overlook others.[185] What is emphasized, what is excluded, and, most importantly, *why* things are emphasized or excluded, is unclear.

---

182. *See supra* note 58 and accompanying text.

183. Though meaningful consent to data collection is at best debatable. *See* Katherine J. Strandburg*, Free Fall: The Online Market's Consumer Preference Disconnect*, 2013 U. CHI. LEGAL F. 95, 133 (2013) (stating "[f]irst, users lack information about the types of harms that may arise from data collection, the prevalence of those harms, and their costs. Second, users lack detailed and useful information about company practices involving data collection, storage, and use. Third, users lack information about how any given instance of data collection fits into the data about them that is already flowing in the online ecosystem. Without these three types of information, Internet users cannot make meaningful assessments of the marginal expected disutility of any given use of an online product or service. Even if they had the necessary information, bounded capacity for information processing and bounded rationality would interfere with their ability to assess their expected disutility and compare it to the expected utility of a given online product or service.").

184. *See* Kevin D. Haggerty & Richard V. Ericson, *The Surveillant Assemblage*, 51 BRIT. J. SOC. 605, 606 (2000).

185. *See* Gillespie, *supra* note 26, at 174 (referring to Ellen Balka and noting that "the slippage between anticipated user [reflected in the shadow bodies] and the user themselves that they represent can be either politically problematic, or politically productive").

These "sufficient[ly] approximat[e]"[186] digital versions of ourselves are compared with sufficiently approximate digital versions of others to determine who "people like us" are and to identify content these "people like us" have engaged with. Some version of that content is what ultimately makes it on our news feeds and supposedly represents what we want to see.

A previous article has developed a complete account of why personalization algorithms undermine individual autonomy and do not represent users' true choice in the selection of content.[187] What this Article now seeks to add, is that this asymmetrical relationship—in terms of information and power—between platforms and their users is exactly what may disqualify them from even limited protections of First Amendment doctrine. If this analysis is correct, regulation of personalization algorithms would not even need to pass the *Central Hudson* test, or any other equivalent test which may be developed for the new category of content navigation algorithms.

The algorithmic knowledge logic of personalized news feeds refers to and produces a type of *public* that is different than the one that is the subject of First Amendment doctrine. Traditional newspapers perform a function in the public sphere by uniting strangers via exposure to common texts.[188] These strangers become a *public*, capable of possessing an opinion and bringing self-government to life.[189] This traditional *public*, brought to life by the printed word, "need not be especially rational, . . . [b]ut it must exist in the imagination of a population."[190] The look-alike publics that personalization algorithms refer to, however, are "calculated publics," that did not exist until the platform's determined their members are *alike*.[191] They are discrete sets of users that are transformed into an audience, and only the platforms know its precise membership.[192] At some level of abstraction, these are strangers that *already* have the same opinions and endorse the same texts: *people like us*. However,

---

186. *Id.* at 174.
187. *See* Grafanaki, *supra* note 73; *see also* Grafanaki, *supra* note 91.
188. *See* Post, *supra* note 56, at 1018.
189. *See id.*
190. *Id.* at 1018–19.
191. Gillespie, *supra* note 26, at 188–189.
192. *Id.* at 189.

unlike the viewers of Fox News and MSNBC, or the readers of the same partisan newspapers, these algorithmic publics may sound intuitive in their description, but are completely opaque in practice; "these algorithmically generated groups may overlap with, be an inexact approximation of, or have nothing whatsoever to do with the publics that the user sought out."[193]

III. Beyond Content Navigation

A. Platform Interface Design: Other Ways to Regulate the Architecture

In the context of platforms, besides code that functions as content curation (whether done through content moderation systems or through content navigation algorithms), there are also other elements of the architecture that affect the conditions of online speech and can be regulated.

These deserve a separate and complete analysis, but let us briefly touch upon one such element in this section: platform interface design.[194] This type of code should be the least problematic to regulate. In the context of First Amendment doctrine, such regulation would resemble content-neutral *time-place-manner* restrictions, which can be perfectly constitutional. An example will help illustrate: If we look at the way the Facebook's News Feed is currently designed, sponsored content appears on a user's feed exactly the same way content from sources the user follows appears, but for the words *suggested post* on the top-left in light grey font. Sometimes the prominent headline is that a user's friend likes a page or a post, but this

---

193. *See id.*

194. *See* Gillespie, *supra* note 26, at 177 (contrasting the early concerns of blending advertisements and organic content with the current concerns that are multidimensional and noting that "the landscape of the Facebook News Feed . . . can no longer be described as two distinct territories, social and commercial; rather, it interweaves the results of algorithmic calculations (what status updates and other activities of friends should be listed in the feed, what links will be recommended to this user, which friends are actively on the site at the moment), structural elements (tools for contributing a status update, commenting on an information element, links to groups and pages), and elements placed there based on a sponsorship relationship (banner ads, apps from third-party sites)").

can either be an organic post or a sponsored post. While the content is labeled, the label itself is almost designed to integrate into the post in a way that goes unnoticed.[195] Platforms are making money from the *sponsored* or *suggested* posts but not from organic posts, and the more clicks or eyeballs the paid content gets, the better for their bottom line.

We can imagine a regulation requiring platforms to clearly separate organic content from sponsored content, not by simply labeling the post, but by having paid content appear at different parts of the screen. This would ensure that readers (the *public*) are not confusing the origins of content and do not, or cannot, share paid content the same way they can share other types of content. The effect of such a regulation could have a major impact for problems like fake news, without even addressing the actual content.

B. Market Forces and Platform Self–Regulation

> "The radio . . . is a mighty force for breaking . . . down [those qualities upon which the enterprise of self-government depends]. . . . And that catastrophe . . . reveals how hollow may be the victories of the freedom of speech when our acceptance of the principle is merely formalistic. Misguided by that formalism we Americans have given to the doctrine merely its negative meaning. We have used it for the protection of private, possessive interests with which it has no concern. It is misinterpretations such as this which, in our use of the radio, the moving picture, the newspaper and other forms of publication, are

---

195. *See* Craig Silverman et al., *In Spite of the Crackdown, Fake News Publishers Are Still Earning Money From Major Ad Networks*, BUZZFEED NEWS (Apr. 4, 2017, 9:05 AM), https://www.buzzfeed.com/craigsilverman /fake-news-real-ads?utm_term=.gggB4Wmw5#.tmj8lAZMV; *see also* Syed, *supra* note 48, at 353.

 giving the name 'freedoms' to the most flagrant enslavements of our minds and wills."[196]

<div align="right">Alexander Meiklejohn, 1948</div>

"Markets are generally good things, both for ordinary products and for speech. But when the legal creation of a market has harmful consequences for free expression - and it sometimes does - then we must reevaluate it in light of free speech principles."[197]

<div align="right">Cass Sunstein, 1990</div>

These passages are quoted to illustrate that the idea of a market failure in the context of markets for information about public affairs is not novel. This is not an area where we can expect the market to correct itself; nor can we rely on the platforms to self-regulate.   The conflicts of interest and information asymmetries are such that we are faced with a market failure. What drives markets and produces welfare is the satisfaction of individual preferences that are exogenous and preexist market relations.[198] This idea of a market, however, is entirely inconsistent with the social and behavioral shaping enabled by the platforms' algorithmic tools.[199]

The economic interests of the platforms dictate keeping their users happy in order to stay on the platform; such incentives are not necessarily aligned with the users' best interests when selecting which content to show them.   For example, fake news may be the kind of content that increases engagement on a platform. As long as users do not protest against such content and consume it, the *economic* incentives of the platforms would dictate to leave it alone. Section 230 allows

---

196.   ALEXANDER MEIKLEJOHN, FREE SPEECH AND ITS RELATION TO SELF-GOVERNMENT 104–05
(Harper & Bros. 1948).

197.   Sunstein*, supra* note 99 at 277.

198.   *See generally* Yochai Benkler & David D. Clark, *Introduction*, 145 DAEDALUS: J. AM. ACAD. ARTS & SCI. 5, 5 (2016).

199.   *See* Yochai Benkler, *Degrees of Freedom, Dimensions of Power*, 145 DAEDALUS: J. AM. ACAD. ARTS & SCI. 18, 23-24 (2016) (noting that, while a critique of markets based on the "endogeneity of preferences" is not a new theme, big data has given it new dimensions).

them to do so without facing liability risks.[200] The economic incentives change only when there is enough public concern about fake news that can influence user engagement. The only reason why platforms would self-regulate is to proactively avoid government regulation. But despite recurring incidents which have caused public outcry, the platforms' proactive measures have proved inadequate every time.[201]

Even in cases where the platforms "mistreat" users, the market may temporarily punish platforms to an extent, as for example happened with the emergence of the #DeleteFacebook movement after the Cambridge Analytica story broke, but such reactions are not effective deterrents for future misbehavior.[202] In fact, the irony of the #DeleteFacebook movement is that its message was spread through the use of hashtags, a filtering tool made available by social media platforms themselves.[203]

Besides the asymmetries of information and power, scholars critique the very idea of "consumer choices" in the context of the communications system.[204] Cass Sunstein has emphasized the distinction between "consumer sovereignty" and "political sovereignty," the former being the idea behind free markets and the latter the idea behind free nations.[205] As consumers, free markets may serve us well; as citizens, however, this is not always the case. Political sovereignty entails the considered judgments of the citizens—the aspirations of the public as a

---

200. *See* 47 U.S.C. § 230 (2012).

201. *See generally* Josh Constine, *Facebook and the Endless String of Worst-Case Scenarios*, TECHCRUNCH (Mar. 18, 2018), https://techcrunch.com/2018/03/18/move-fast-and-fake-things/.

202. *See generally* Jessica Guynn, *Delete Facebook? It's a Lot More Complicated Than That*, USA TODAY (Mar. 28, 2018, 3:59 PM), https://www.usatoday.com/story/tech/news/2018/03/28/people-really-deleting-their-facebook-accounts-its-complicated/464109002/.

203. *See generally* Sandra González-Bailón, *Want to Change Facebook? Don't Delete Your Account—Use It for Good*, QUARTZ (Apr. 4, 2018), https://qz.com/1244750/the-delete-facebook-movement-is-ultimately-self-defeating/.

204. *See* Sunstein, *supra* note 99, at 287; *see also* SUNSTEIN, *supra* note 30, at 52–57.

205. SUNSTEIN, *supra* note 30, at 52–54 (stating "the notion of consumer sovereignty underlies enthusiasm for the Daily me [and] . . . is the underpinning of any utopian vision of the unlimited power to filter. . . . The notion of political sovereignty underlines the democratic alternative.").

2018   *REGULATING THE FILTERS*   161

whole[206]—and, unlike consumer sovereignty, which views individuals as *having* fixed tastes and preferences there to be discovered, political sovereignty understands individuals and communities as being *shaped* by the political process. In the context of platforms, the two concepts can get especially blurred. The same platform appeals to both our role as consumers, expecting to choose exactly as we wish, and to our role as citizens of a democratic society, requiring information about public affairs.

Conclusion

We cannot respond to the recent fake news crisis by making fake news illegal or by holding platforms liable for hosting fake news like the Malaysian government recently did.[207] However, the problem was never in the existence of fake news, a phenomenon that has always resided in the fringes of the information ecosystem. The recent crisis was a result of the current ecosystem's architecture, which is responsible for the way information gets amplified and becomes more persuasive. The assertion put forward in this Article is that we should regulate certain elements of the underlying architecture, that we can do so without doctrinal barriers, and that doing so can address some of the problems that the new ecosystem has made more salient.

Today's information environment has changed the conditions of speech. The new conditions are not aligned with the prerequisites for citizens in a democratic society to perform their civic duties, and they undermine the notion of true autonomous choice in the selection and consumption of content. It was therefore argued that certain algorithms used by social media platforms to filter content can and should be regulated.

This argument was primarily structured on a distinction between algorithms that filter content for hosting purposes versus algorithms that filter content for navigation purposes.

---

206. *Id.* at 54 (stating "political sovereignty embodies democratic self-government, understood as a requirement of 'government by discussion,' accompanied by reason-giving in the public domain, where different people speak with one another and listen respectfully, even when in intense conflict").

207.   *See supra* note 47 and accompanying text.

162   *PACE LAW REVIEW*   Vol. 39.1

This distinction can help categorize the growing body of literature and, most importantly, can describe the limits of the First Amendment as a deregulatory tool in this context. Doctrinal barriers preempt regulation of platform content moderation algorithms, meaning algorithms that implement the policies of platforms for deciding to host or censor content items. Algorithms that implement content moderation policies are the closest to an editorial analogy and, arguably, deserving of full First Amendment protection.

On the other hand, the challenges are not as strong when it comes to how platforms facilitate content navigation. This Article firstly argued that content navigation algorithms should not be viewed as speech for First Amendment purposes. Further, it was argued that even if we were to view content navigating algorithms as speech, this should not be political speech that is subject to core First Amendment doctrine. Rather, it should be considered a new category of protected speech that falls outside the public discourse. Much like commercial speech, it is less problematic to regulate. Even further, this Article advanced an argument limited to personalization algorithms in particular, suggesting that these may not even deserve the limited protections of the type available to commercial speech, as they are better categorized as commercial communications that fall outside the protected speech categories altogether.

Designing a regulatory intervention that can prove effective is a complex and challenging task that involves future interdisciplinary work. What this paper demonstrates is that the First Amendment is not a barrier to taking on such a task.