THE AUTOMATED PUBLIC SPHERE

ROSALUXEMBURG STIFTUNG
NEW YORK OFFICE

By Frank Pasquale
Table of Contents

Pathologies of the Automated Public Sphere. By the editors........................................................1

The Automated Public Sphere

By Frank Pasquale...............................................................................................................................2

The Long Shadow of Intermediary Irresponsibility..............................................................3

Toward a Robust Regime of Intermediary Responsibility.......................................................5

Concerns and Concessions........................................................................................................9

Conclusion: A Return to Professionalism................................................................................12

References...................................................................................................................................13

Published by the Rosa Luxemburg Stiftung, New York Office, December 2017

Editors: Stefanie Ehmsen and Albert Scharenberg
Address: 275 Madison Avenue, Suite 2114, New York, NY 10016
Email: info.nyc@rosalux.org; Phone: +1 (917) 409-1040

With support from the German Foreign Office.

The Rosa Luxemburg Foundation is an internationally operating, progressive non-profit institution for civic education. In cooperation with many organizations around the globe, it works on democratic and social participation, empowerment of disadvantaged groups, alternatives for economic and social development, and peaceful conflict resolution.

The New York Office serves two major tasks: to work around issues concerning the United Nations and to engage in dialogue with North American progressives in universities, unions, social movements, and politics.

www.rosalux-nyc.org
Pathologies of the Automated Public Sphere

The ubiquity of the internet today—namely of social media networks and large search engines—has complicated the ways in which content is produced and received, deeply altering how society thinks about the rights to free speech, freedom of the press, and freedom of expression. Public speech can now easily be spread worldwide while remaining anonymous. Nazi-related content liable to prosecution in Europe is safely hosted on US servers where it is protected by the First Amendment. On top of this, our access to content has become completely mediated by algorithms designed to maximize profits. This transformation in the production, distribution, and consumption of content has become the inexorable backdrop of contemporary debates on the basic right to freedom of speech.

As automation emerges as a problem affecting all spheres of production, we are increasingly confronted with its ethical implications. This is especially true for the discussions spurred by the new digital contexts that shape public opinion. The automation of decision-making processes put in motion by digital platforms in sensitive areas such as editing, moderating, advertising, and circulating information is at the source of many controversies. When the ways in which information and opinion are produced and disseminated become open to manipulation, we are forced to deal with the consequences—an unregulated platform that takes advantage of the same capitalist logic that undermines society in so many other ways.

In this new piece, Frank Pasquale, affiliate fellow at Yale Law School's Information Society Project and author of *The Black Box Society: The Secret Algorithms That Control Money and Information*, argues that powerful interest groups build their dominance with the help of a regulatory regime that permits secrecy and complexity. The study tackles events surrounding the recent US elections as well as other cases in which online interventions (or lack thereof) have allowed for the spread of hateful ideologies in the broader public. Presenting a series of legal and educational steps, the author shows how we may curtail the effects of the pathologies that the contemporary automated public sphere creates.

Disrupting secretive business strategies, preventing the concentration of power and money into a few hands, and introducing more instances of human intervention are initiatives that put the focus on regulating the power held by big tech companies. However, other proposed measures trigger crucial ethical considerations around this regulatory apparatus. How can we ensure that regulation does not turn into censorship, suppression, and thereby into another tool for manipulation and control of society?

Beyond fully endorsing these proposals or necessarily opposing all such forms of regulation, we believe that as progressive actors—many times on the losing end of digital control and/or harassment—we need to reconsider our strategies and methods under new contexts. How can we rethink regulation to make it work in both fair and transparent ways everywhere? Can we devise ways to regulate users, who play an important part in producing and distributing content, without falling into brute censorship? Are these desirable, forward-looking options, or rather desperate reactions to the reality of today’s digital world? Ultimately, we need to ask, what is the role of the internet in society today and how can we improve the digital environment for all of us?

*Stefanie Ehmsen and Albert Scharenberg*

*Co-Directors of New York Office, December 2017*
The Automated Public Sphere

By Frank Pasquale

As internet usage grew in the 2000s, scholars promoted its emancipatory potential. Yochai Benkler praised not only the wealth that would be promoted by networks, but also its distribution—toward a platform of platforms that would enable millions of new voices to be heard online (Benkler, 2007). This optimism also animated one of the United States’ seminal cases on internet regulation, *Reno v. ACLU* (1997), which presumed the openness of the internet would redound to the benefit of all. The majority opinion in ACLU darkly cautioned the US government to avoid mucking about in many forms of internet regulation, lest it infringe on free expression rights in an online environment that the justices, as well as later boosters, idealized. Large platforms themselves harbor utopian pretensions to this day; for example, Mark Zuckerberg has marketed Facebook as a nascent global community (even as social critics lament how time online diverts citizens from in-person engagement with friends and neighbours) (Rushkoff, 2016).

Even in the 1990s, scholars warned about the implications of deregulating the internet (Chin, 1997). By the mid-2010s, it is hard to remain optimistic about the role of the internet in organizing a new, and critically important, digital public sphere. Wealth has emerged in online advertising, but it is largely claimed by two firms—Google and Facebook take about 75% of the $73 billion digital advertising market in the US (Bond, 2017). These information intermediaries are driven by profit, and their methods of selecting and arranging newsfeeds and search engine results pages are secret (Pasquale, 2015b, pp. 59-100). The promised *Wealth of Networks* has given way to a black box society—one where trolls, bots, and even foreign governments ma-raud to distort the information environment on Twitter, Facebook, Google News, Reddit, and other networks.

We now know that virtually every positive promise made about the internet in the early 2000s has a shadow side. While secrecy has empowered some voices who would otherwise be afraid to speak up, it has also protected trolls, doxers, and other bad actors online who silence others’ speech via intimidation. Moreover, online anonymity is of a piece with financial anonymity, which has empowered thousands of shell companies to obscure who is actually funding messages that could sway the public, legislators, and regulators. Everyone is invited to participate, but so too is “everyone” capable of disrupting other communities of interest online, via hashtag spamming or trolling—whether by civil society groups, state actors, or miscreants pursuing disruption “for the lulz.” First celebrated as a way to hold states accountable for illegal actions, Wikileaks has emerged as a witting agent of authoritarian state interference in elections with a troubling tendency to emit anti-Semitic messages. While major content owners have found their grip on public attention diminished, fragmentation of audiences has given megaplatforms unprecedented global power over attention-commanding interfaces.

That last reversal is the subject of this essay. In a world of stable and dominant media firms, large social networks and search engines were in a rough equilibrium of power relative to the owners and creators of the content they selected and arranged (Pasquale, 2010). However, a general trend toward media revenue decline (and platform revenue growth) makes a new endgame
apparent: online intermediaries as digital bottle-necks or choke-points, with ever more power over the type and quality of news and non-news media that reaches individuals (Bracha and Pasquale, 2008; Pasquale, 2008b). The source of this power is, at bottom, big data—the ability of megaplatforms to accumulate ever-more-intimate profiles of users, which are then of enormous interest to commercial entities, political campaigns, governments—indeed, anyone aspiring to monitor, monetize, control, and predict human behavior.

Large online intermediaries tend to reduce at least one good type of media pluralism, and to promote a very destructive type of diversity. They make the metric of success online “virality,” promoting material that has received a good deal of attention or seems to match a sub-public’s personalization profile, regardless of whether it is true or minimally decent (Pasquale, 2006). That reduces pluralism by elevating profit considerations over the democratizing functions of public discourse, and effectively automating the public sphere. Decisions that once were made by humans are now made by profit-maximizing algorithms. Moreover, the same intermediaries also promote a very troubling diversity by permitting themselves to be manipulated by the most baseless and dangerous propagandists (Marwick and Lewis, 2017). Such political forces are particularly gifted at creating media capable of influencing and persuading low-information, floating voters—exactly the persons most likely to swing the results of elections.

This essay first describes the documented, negative effects of online propagandists’ interventions (and platforms’ neglect) in both electoral politics and the broader public sphere (Part I). It then proposes several legal and educational tactics to mitigate their power, or to encourage or require them to exercise it responsibly (Part II). The penultimate section (Part III) offers a concession to those suspicious of governmental intervention in the public sphere: some regimes are already too authoritarian and generally unreliable to be trusted with extensive powers of regulation over media (whether old or new media), or intermediaries. However, I conclude that the inadvisability of extensive media regulation in disordered societies only makes this agenda more urgent in well-ordered societies, lest predictable pathologies of the automated public sphere accelerate the degradation of their democracies.

The Long Shadow of Intermediary Irresponsibility

Jürgen Habermas observed in 1962, “the process in which societal power is transformed into political power is as much in need of criticism and control as the legitimate exercise of political domination over society” (Habermas, 1962, trans. 1991, p. 210). As part of the Frankfurt School, Habermas was sensitive to the ways in which new media, manipulated by both corporate and state actors, had fundamental-
ly altered processes of democratic will formation. He deemed such transitions a “structural transformation” of the public sphere, since new forms of media both accelerated, altered, and reshaped the expression critical both to opinion formation and expression.

The basic contours of mass media-driven politics and culture remained stable through much of the second half of the twentieth century. But since the mid-1990s, the public sphere has endured yet another structural transformation. Megafirms like Facebook and Google have largely automated the types of decisions once made by managers and programmers at television networks, or editors at newspapers. Automated recommendations are often helpful, aiding audiences as they seek to sort out the blooming, buzzing confusion of topics online. But they are also destabilizing traditional media institutions and circuits of knowledge.

For example, the US election featured deeply disturbing stories about manipulation of social media for political ends. Unreliable sources proliferated, particularly among right-wing echo chambers. In December 2016, a Facebook-fueled fake news story about Hillary Clinton prompted a man to arrive in a pizza parlor in Washington D.C. with a gun (Abrams, 2016). The fake story reportedly originated in a white supremacist’s tweet. Politically motivated, profit-seeking, and simply reckless purveyors of untruths all prospered. A Macedonian teen churned out stories with no basis, tarring Hillary Clinton with an endless series of lies, in order to score quick profits (Smith and Banic, 2016). For profit-minded content generators, the only truth of Facebook is clicks and ad payments. Bence Kollanyi, Phil Howard, and Samuel Woolley estimated that tens of thousands of the tweets “written” during the second US presidential debate were spewed by bots (Kollanyi, Howard, and Woolley, 2016). These bots serve multiple functions—they can promote fake news, and when enough of them retweet one another, they can occupy the top slots in response to tweets from candidates. They can also flood hashtags, making it very difficult for ad hoc publics to crystallize around an issue.

On Facebook, a metastatic array of fake content generators and hard-core partisan sites flooded news feeds with lies and propaganda. Facebook, as usual, disclaimed any responsibility for the spread of stories falsely claiming that the Pope had endorsed Donald Trump, or that Hillary Clinton is a satanist (to give a mild pair of the lies that swarmed the platform) (Schaedel, 2016; Evon, 2016). But the Silicon Valley firm has several levels of responsibility.

Basic design choices mean that stories shared on Facebook (as well as presented by Google’s AMP) all look very similar, and indeed have for years (Chayka, 2016). Thus a story from the fabricated “Denver Guardian” can appear as authoritative as a Pulitzer Prize-winning *New York Times* investigation (Lubbers, 2016). More directly, Facebook profits from fake news—the more a story is shared (whatever its merits), the more ad revenue it brings in (Molina, 2016). Most disturbingly, we now know that Facebook directly helped the Trump campaign target its voter suppression efforts at African-Americans (Winston, 2016).

Google has suffered from other racially tinged scandals (Noble, 2018). Consider, for instance, recurrent problems with Google’s “autocompletes”—when Google anticipates the rest of a search query from its first word or two. Google autocompletes have often embodied racist and sexist stereotypes (Cadwalladr, 2016). Its image search has also generated biased results, absurdly and insultingly tagging some photos of black people as “gorillas” (Guarino, 2016; Barr, 2015).

If Google and Facebook had clear and publicly acknowledged ideological agendas, users could grasp them and “inoculate” themselves
accompanying, with skepticism toward self-serving views (Pasquale, 2011). However, the platforms are better understood as tools rapidly manipulated to the advantage of search engine optimizers, well-organized extremists, and others at the fringes of political respectability or scientific validity. Thus a search for “Hillary’s Health” in October 2016 would have led to multiple misleading videos and articles groundlessly proclaiming that the US Democratic presidential candidate had Parkinson’s Disease. Google search results reportedly helped shape the racism of Dylann Roof, who murdered nine people in a historically black South Carolina church in the US in 2015. Roof said that when he googled “black on white crime,” the first website he came to was the Council of Conservative Citizens,” which is a white supremacist organization. “I have never been the same since that day,” he said. So too are sources of support for climate denialsists, misogynists, ethnonationalists, and terrorists easily developed and cultivated in what has become an automated public sphere.

These terrifying acts of violence and hate are likely to continue if action is not taken. Nor is democracy safe in a carelessly automated public sphere. Without a transparent curation process, the public has a hard time judging the legitimacy of online sources. In response, a growing movement of academics, journalists, and technologists is calling for more algorithmic accountability from Silicon Valley giants (Pasquale, 2015a). As algorithms take on more importance in all walks of life, they are increasingly a concern of lawmakers. And there are many steps both Silicon Valley companies and legislators should take to move toward more transparency and accountability.

Toward a Robust Regime of Intermediary Responsibility

Activist and academic responses to these imbroglios have been multifaceted. Some communication scholars have rightly criticized Facebook for its apparent indifference to the problem of fake or misleading viral content (Tufekci, 2016). Others have focused their ire on the mainstream media, claiming that it was the recklessness or lack of professional responsibility at right-wing news sources (and established media institutions like CNN and The New York Times) which accelerated the rise of authoritarian candidates like Trump (Kreiss, 2016; Robinson, 2016).

In truth, there is no contradiction between a critique of the new media and deep disappointment in old media. Moreover, any enduring solution to the problem will require cooperation between journalists and coders. Facebook can no longer credibly describe itself as merely a platform for others’ content, especially when it is profiting from micro-targeted ads (Pasquale, 2016a). It has to take editorial responsibility. So, too, should mega-platforms like Google take on some basic responsibilities for the content they distribute. This section describes several specific initiatives that would help counter the discrimination, bias, and propaganda now too often polluting (and even overwhelming) online spaces.

A. Label, monitor, and explain hate-driven search results.

In 2004, anti-Semites boosted a Holocaust-denial site called “Jewwatch” into the top 10 results for the query “Jew” (Pasquale, 2006). Ironically, some of those horrified by the site may have
helped by linking to it in order to criticize it. The more a site is linked to, the more prominence Google's algorithm gives it in search results. The Anti-Defamation League and other civil rights organizations complained to Google about its technology's gift of prominence to entities eminently unworthy of such attention.

Google responded to complaints by adding a headline at the top of the page entitled “An explanation of our search results.” A web page linked to the headline explained why the offensive site appeared so high in the relevant rankings, thereby distancing Google from the results. The label, however, no longer appears. In Europe and many other countries, lawmakers should consider requiring such labeling in the case of obvious hate speech. To avoid mainstreaming extremism, labels may link to accounts of the history and purpose of groups with innocuous names like “Council of Conservative Citizens” (Pasquale, 2016a; Pasquale, 2008a).

Are there free expression concerns here? Not really. Better labeling practices for food and drugs have escaped First Amendment scrutiny in the US, and why should information itself be different? As law professor Mark Patterson has demonstrated, many of our most important sites of commerce are markets for information: search engines are not offering products and services themselves but information about products and services, which may well be decisive in determining which firms and groups fail and which succeed (Patterson, 2017). If they go unregulated, easily manipulated by whoever can afford the best search engine optimization, people may be left at the mercy of unreliable and biased sources.

B. Audit logs of the data fed into algorithmic systems.

We should expect any company aspiring to order vast amounts of information to try to keep its methods secret, if only to reduce controversy and foil copycat competitors. However wise this secrecy may be as a business strategy, it devastates our ability to truly understand the social world Silicon Valley is creating. Moreover, like a modern-day Ring of Gyges, opacity creates ample opportunities to hide anti-competitive, discriminatory, or simply careless conduct behind a veil of technical inscrutability.

A recurring pattern has developed: some entity complains about a major internet company’s practices, the company claims that its critics don’t understand how its algorithms sort and rank content, and befuddled onlookers are left to sift through rival stories in the press. Massive search operations are so complex, and so protected by both legal and de facto secrecy, that it is almost always impossible for those outside a search engine or social network firm to identify all the signals that are driving a given set of results. Silicon Valley journalists tend to give their advertisers the benefit of the doubt; national media outlets find the mysteries of online content ordering perfectly fit into their own templates of balanced reporting. No one knows exactly what’s going on when a dispute arises, so rival accounts balance into an “objective” equipoise.

Regulators need to be able to understand how some racist or anti-Semitic groups and individuals are manipulating search and social media feeds (Pasquale, 2010). We should require immutable audit logs of the data fed into algorithmic systems. Machine-learning, predictive analytics, or algorithms may be too complex for a person to understand, but the data records are not. They can be subjected to algorithmic audits.

A relatively simple set of reforms could vastly increase the ability of entities outside Google and Facebook to determine whether and how the firms’ results and news feeds are being manipulated. There is rarely adequate profit motive
for firms themselves to do this—but motivated non-governmental organizations can help them be better guardians of the public sphere.

C. Ban certain content.

In cases where computational reasoning behind search results really is too complex to be understood in conventional narratives or equations intelligible to humans, there is another regulatory approach available: to limit the types of information that can be provided.

Though such an approach would raise constitutional objections in the US, nations like France and Germany have outright banned certain Nazi sites and memorabilia. Policymakers should also closely study laws regarding “incitement to genocide” to develop guidelines for censoring hate speech with a clear and present danger of causing systematic slaughter or violence against vulnerable groups. It’s a small price to pay for a public sphere less warped by hatred. And unless something like it is done, expect social media-driven panics about despised minorities to lead to violence.

To be sure, this approach would almost certainly draw immediate legal action in the United States, where a form of “free expression fundamentalism” has protected even the most reprehensible speech. Cyberlibertarians support First Amendment protections for algorithmic orderings of information. Relatedly, the same scholars and judges eager to protect the “speech” of computers also promote the idea that massive corporations’ “expression” is deserving of exceptional protection from the very state so often suborned or coopted by those same corporations.

This science fictional appeal to Asimov-ian ideals of “speaking robots” has fed a romanticization of corporate speech. The logical endpoint is a continual “battle for mindshare” by various robot armies, with the likely winner being the firms with the funds to hire the top programmers and the network effect dynamics to gather the most data for the optimal crafting of messages for microtargeted populations. It goes without saying that this type of decomposition of the public sphere does not represent a triumph of classic values of free expression (autonomy and democratic self-rule); indeed, it portends their evaporation into the manufactured consent of a phantom public.

D. Permit outside annotations and hire more humans to judge complaints.

In the US and elsewhere, limited annotations—“rights of reply”—could be permitted in certain instances of defamation of individuals or groups (Pasquale, 2008a). Google continues to maintain that it doesn’t want human judgment blurring the autonomy of its algorithms. But even spelling suggestions depend on human judgment, and in fact, Google developed that feature not only by means of algorithms but also through a painstaking, iterative interplay between computer science experts and human beta testers who report on their satisfaction with various results configurations. As Sarah Roberts, Lily Irani, and Paško Bilić have shown, supposedly digitized companies are constantly reliant on manual interventions by human beings (Bilić, 2016; Irani, 2013; Roberts, 2016a; 2016b). Requiring a few more is not a major burden for these firms.

This step is important because we now know (if we ever doubted) that the hoary “marketplace of ideas” metaphor is misleading. The best ideas are not necessarily the most highly valued; the most sensational or gratifying propaganda can beat out careful reporting. Highly motivated, well-resourced groups can easily manipulate newsfeeds or search engine result pages (SERPs). “Dark ads” and sophisticated personalization algorithms enable con-
stant experimentation on unwitting human research subjects, so A/B testing can reveal exactly what manipulation works best. Without conscientious and professional curation of such algorithmic orderings of information, the public sphere's automation is susceptible to distortion by the most well-resourced and committed entities.

The European Union's commitments to rights to be forgotten, and rights of erasure, show that the algorithmic ordering of information can be a socially inflected process, with fairer modes of participation for citizens and civil society (Pasquale, 2016b).

E. Limit the predation possible by online intermediaries.

Personalization is leading advertisers to abandon traditional, and even not-so-traditional, publishers in favour of the huge internet platforms. No other rival can approach either the granularity or the comprehensiveness of their data. The result is a revolution-in-process about who can afford to keep publishing, and concomitant alarm about the concentration of media clout into fewer and fewer hands.

One platform owner, Jeff Bezos, accumulated wealth equivalent to one hundred times the total value of the United States' second most important newspaper, The Washington Post. He bought the Post, with concomitant chilling effects on the paper's ability to criticize his own business empire-building, or similar strategies by platform capitalists. Given the leverage potential of their own ever-higher expected earnings, large platforms may soon be able to move to buy more content producers themselves, as cable networks and ISPs have done—or perhaps purchase entire cable networks and ISPs. This type of vertical integration would be a major threat to the autonomy of journalism. Given all the negative externalities generated by online intermediaries, policymakers should limit the profits they make relative to revenues of the content owners whose work they depend on. In the health care context in the US, private insurers can only keep a certain percentage of premiums (usually 15 to 20%)—the rest must go to health care providers, like hospitals, doctors, and pharmaceutical firms. Such a rule keeps the intermediary from taking too much of the spending in a sector—a clear and present danger in monopolistic internet contexts, as well. Governments could limit the amount of profits that search engines and social networks make as intermediaries, requiring them to pay some share of their revenues to content generators like newspapers and media firms (Lanier, 2013; Lehdonvirta, 2017). Alternatively, policymakers could simply force large platforms to pay a fair share of the tax they now avoid by shuttling income to tax havens, and use some of that revenue for public broadcasting alternatives.

F. Obscure content that is damaging and not of public interest.

When it comes to search queries on an individual person's name, many countries have aggressively forced Google to be more careful in how it assembles data dossiers presented as SERPs. Thanks to the Court of Justice of the European Union, Europeans can now request the removal of certain search results revealing information that is "inadequate, irrelevant, no longer relevant or excessive," unless there is a greater public interest in being able to find the information via a search on the name of the data subject (Pasquale, 2016b).

Such removals represent a middle ground between information anarchy and censorship. They neither disappear information from the internet (it can be found at the original source, and in searches on terms other than the complaining party's name), nor allow it to dominate the impression of the aggrieved individual. They
are a kind of obscurity that lets ordinary individuals avoid having a single incident indefinitely dominate search results on his or her name. For example, a woman whose husband was murdered 20 years ago successfully forced Google to take news of the murder off search results on her name. This type of public responsibility is a first step toward making search results and social network newsfeeds reflect public values and privacy rights.

Concerns and Concessions

There will be fierce opposition to virtually all of the proposals I have listed above. Some will arise merely out of commercial motivations: policing hate speech and fake news is more expensive than letting it flourish. Platforms would rather just pile up advertising revenue. As Jodi Dean has demonstrated, outrageous content stokes at least as much engagement online as it has in the traditional media (Dean, 2010). Indeed, the problem is easily intensified online, as personalization allows platforms to deliver material precisely targeted to maximize clicks, likes, and shares (Citron, 2014). Slowing that accelerated engagement costs a platform potential advertising, and all-important data about its users (Srnicek, 2017). It also impedes the platform's ability to shape its users into the kind of people who uncritically act in behaviorally manipulable ways (Schüll, 2012).

Unless platforms can demonstrate that the intermediary responsibilities discussed above would compromise their ability to run the platform at a reasonable rate of return, such cost-based objections should be dismissed. Neither Mark Zuckerberg nor Facebook shareholders have any legitimate expectation of permanent, massive returns on their investment. Indeed, impeding their ability to accumulate the surplus they have used to buy rival and adjacent firms may well encourage innovation (Stucke and Grunes, 2017).

Many apologists for big tech firms claim that this type of responsibility is impossible (or unwise) for a firm like Facebook to take on (Turton, 2016; Lessin, 2016). They argue that the volume of shared content is simply too high to be managed by any individual, or team of individuals. But this argument ignores the reality of continual algorithmic and manual manipulation of search results at Google. As technology writer Timothy Lee explains,

During the 2000s, people got better and better at gaming Google's search algorithm. Some were running quasi-media companies whose writers churned out dozens of extremely short, poorly researched articles based on popular search terms. [...] In a January 2011 blog post, Google search quality czar Matt Cutts acknowledged that Google had a big problem with these “content farms.” [...] Later that year, Google brought down the hammer, releasing changes to its search algorithm that caused traffic at major content farms to plummet.

[This] represented Google making a deliberate value judgment that some kinds of content were worse than other kinds. Early versions of Google took a naively data-driven approach, assuming that a link from one site to another was a sign of quality. [...] [In later, more sophisticated iterations.] Google include[d] human reviewers in the mix because algorithms inevitably make mistakes and manual human review is needed to keep the algorithms on the right track. Previously reviewed pages can be fed back into Google's software, allowing the algorithms to learn from human judgment and get better over time. So Facebook doesn't have to choose between fighting fake news with algorithms or human editors. An effective fight against fake news is going to require heavy use of both approaches (Lee, 2016).
There are powerful lessons in this passage. First, be wary of the convenient self-reification of platforms. Facebook may aspire to be merely a technology company. Those aspirations may express themselves as a petulant insistence that unsupervised, rather than supervised, machine learning is the ideal way to solve problems on the platform. But that “identity” is a constructed and convenient one, directly at odds with tech firms’ repeated invocation of free expression protections to shield their actions from governmental scrutiny (Pasquale, 2016c).

Beyond economic and technical objections, there is a third, deeper objection to intermediary responsibility, focusing on the regulatory apparatus necessary to make it meaningful and robust. Authoritarian regimes have tried to stifle political dissent by regulating Facebook and Google. For example, the Thai, Russian, Chinese, and Turkish governments have aggressively policed criticism of national leaders, and have intimidated dissidents. Corrupt governments may be susceptible to excessive influence from well-organized lobbies. Fossil fuel lobbyists may influence regulators to force intermediaries to monitor and censor environmental activists committed to resistance against pipeline projects (Citron and Pasquale, 2011, p. 1445; ACLU, 2017). Overly onerous annotation requirements, or rights to be forgotten, may become a pretext for driving a popular platform out of a country. Governments may abuse taxation powers, too, in retaliation against a platform that enables stinging or politically effective criticism of them. Or platforms may successfully lobby to have their own personnel and allies appointed to the agencies and commissions set to regulate them. A search or robotics or social network commission, for example, might start out with a robust agenda, but over years or decades may find itself taken over by appointees closely aligned with dominant industry players.4 Still, there is little reason to assume that the actions of the worst governments are likely in other, more developed and democratic public spheres. Indeed, intervention in the public sphere while a polity is still well-ordered may be the only way to keep it well-ordered. Some of these concerns are variations on the classic problem of regulatory capture: the very institutions meant to regulate an industry may be taken over by that industry. Fortunately, the problem has now been so carefully studied that many prophylactic measures could be put in place to avoid it (Carpenter and Moss, 2014). Revolving door rules could prevent officials and bureaucrats from working for the industry they are regulating for five or ten years after they depart their agency. Higher pay for regulators would also help assure more independence, as would the type of automatic funding mechanism that now empowers the United States Consumer Financial Protection Bureau to act as that country’s lead consumer-oriented financial regulator. While serious, the problem of regulatory capture is not insurmountable.

More serious is a larger problem of circularity, well-identified by Charles Lindblom: the ability of powerful economic entities to take over political institutions and use that political power to enhance their economic power, which in turn gives them resources necessary to entrench their political power (Lindblom, 1977, p. 201-213). The rise of oligarchical power in nations around the world suggests how deep the problem of circularity can be (Winters, 2011). The tendency of oligarchs to enact programs that simultaneously harm the material conditions of their electoral base, while cultivating and consolidating their sense of political identity organized around common grievance, should also serve as a spur to reconsider the foundations of the critiques that motivated the program of reform developed above.

For example, consider the classic problem of the filter bubble (Pariser, 2011; Sunstein, 4 Bracha and Pasquale (2008) included the first discussion in legal literature of a regulatory body targeted at search engines.
2007). Personalization enables internet users to ignore points of view they disagree with, so the filter bubble model states and therefore increases polarization. Let us assume, for now, that there is some extant middle ground of consensus worth saving. Extant solutions to the filter bubble dynamic presume, first, that “all sides” or “both sides” can be exposed to some critical mass of opposing or diverse viewpoints via, say, must-carry rules, or some privately implemented version of them (Pasquale, 2016a, p. 499-500). To make that reform tractable, assume for now a binarily divided society, divided between left and right voters. The great problem for advocates of “filter bubble” reforms is that they cannot adequately model whether exposure of one side’s adherents to the other side’s version of facts, priorities, ideology, or values, will lead to understanding or revulsion, reconsideration, or recalcitrance.

To be sure, effects studies in media have been contested for decades. It may be impossible for today’s digital deliberative democrats to demonstrate the empirical likelihood of open-mindedness among voters (Gutmann and Thompson, 2004). But they should be open to understanding the danger of plausible models of asymmetrical openness to opposing views. A society may have a “hard left” and a “soft right,” such that those on the right are quite willing to assess and even adopt some left proposals, while the vast majority of the left is unalterably opposed to accepting any right ideas. In such a scenario, an assault on the filter bubble is only likely to chip away at conservative self-identification among the “soft right,” while succoring the hard left. Perhaps intuiting that danger to its coherence and ability to project power, today’s right in the United States may be inoculating itself against such ideological slippage. Very often, those in the center right will defend or applaud those to their right, but the comity rarely goes the other way (Nagle, 2017).

In a situation of asymmetrical persuadability, filter bubble-inspired reforms will tend only to consolidate the power of the social group or political party most steadfastly committed to maintaining its own position. We can, of course, imagine 12 Angry Men-type scenarios where a small remnant of deeply moral hold-outs uses its reform-granted exposure to others to gradually convince the rest of society of the wisdom of its position. However, just as likely is a splitting of society into the more contemplative and the more active, à la the famed quote on the “reality-based community” from a member of the George W. Bush administration.5

This elementary challenge to filter bubble-driven reform suggests a larger problem with the deliberativist political theory driving reforms of the automated public sphere (Pasquale, 2008c). How can democracy operate when large swaths of the population subscribe to diametrically opposed conceptions of the nature of politics? Consider the deliberativist approach as one end of a spectrum of theories of politics, with a Schmittian, decisionist approach on the opposite end. Deliberativists see politics as fundamentally a realm of reasoned disagreement, culminating in some form of agreement (or at least improved understanding) after debate (Parkinson and Mansbridge, 2012). Jürgen Habermas detailed the “ideal speech situation” as the regulative ideal of such political deliberation, where everyone would either be able to voice their own views, and learn from others, or at least count on their political representatives

---

5 The journalist Ron Suskind authored an article that quoted a senior George W. Bush administration official as saying “that guys like me were ‘in what we call the reality-based community,’ which he defined as people who ‘believe that solutions emerge from your judicious study of discernible reality.’ [...] ‘That’s not the way the world really works anymore,’ he continued. ‘We’re an empire now, and when we act, we create our own reality. And while you’re studying that reality—judiciously, as you will—we’ll act again, creating other new realities, which you can study too, and that’s how things will sort out. We’re history’s actors [...] and you, all of you, will be left to just study what we do.” Suskind (2004).
in a legislative body engaging in a similar process (Habermas, 1991).

Habermas’s conception of parliamentary democracy was part of a long scholarly campaign to lay to rest the type of post-rational, emotivist politics associated with Carl Schmitt (Müller-Doohm, 2017). But Schmitt’s critical ideas are finding more traction today, both in diagnoses of political polarization, and in the actual attitudes and actions of many voters and politicians. For those committed to a Schmittian perspective, there are friends and enemies in politics, and almost no new information can dissuade them from their attachment to their party or leader. US President Donald J. Trump memorably bragged that he could “shoot someone on Fifth Avenue,” and his voters would still remain devoted to him. That is a Schmittian devotion par excellence. More strategically, a political party may change voting rules to entrench its power, creating a self-reinforcing dynamic: the more the rules change in its favor, the more opportunities it has to entrench majorities and super-majorities that enable further rule changes (Daley, 2015; Berman, 2016). In such circumstances, some or all of the reforms mentioned above could backfire, simply adding to the power of a dominant party in a disordered polity, rather than preserving and promoting the type of pluralism that is a hallmark of a well-ordered democracy.

Conclusion: A Return to Professionalism

Given the potential pitfalls of regulating the automated public sphere, implementation of the reform ideas in Part II above should be undertaken with care in well-ordered polities, and may be impossible or counterproductive in disordered polities. But regardless of those difficult distinctions, those in media can do much to respond to the automated public sphere’s infirmities.

Journalists should be more assertive about their own professional prerogatives and identity. In the aftermath of the fake news scandals, Tim O’Reilly asserted that decisions about the organization of newsfeeds and presentation of information in them were inherently algorithmic functions, to be supervised by the engineers at Facebook (O’Reilly, 2016). Certainly the alpha geeks whom O’Reilly describes as his subject share that view: the human editors of trending topics at Facebook were low status contract workers, who were unceremoniously dumped when a thinly sourced news story asserted that conservative content was being suppressed (Ohlheiser, 2016; CBS News, 2016). Shortly thereafter, Facebook was swamped by the fake news which now is the topic of so much controversy. The real lesson here is that human editors at Facebook should be restored, should be given more authority, not less, and that their deliberations should be open to some forms of scrutiny and accountability.

Some communication scholars have resisted the idea of professionalization of online content creation, curation, and delivery in the name of citizen journalism which would democratize the power of the press to anyone with a computer and an internet connection. While a beautiful idea in theory, in practice a failure among the de facto sovereigns of the internet to distinguish between stories on the real Guardian and the “Denver Guardian” is not simply a neutral decision to level the informational playing field. Rather, it predictably accelerates propaganda tactics honed by millions of dollars of investment in both data brokerages and shadowy quasi-state
actors now investigated by the CIA as sources of bias, disinformation, and illegal influence in the election (Revesz, 2016; Feldman, 2016). Freedom for the pike is death for the minnows.

In the 1980s, the chair of the US Federal Communications Commission, Mark Fowler, dismissed the bulk of regulation of broadcasters as irrelevant, since he viewed the television as nothing more than “a toaster with pictures” (Boyer, 1987). In the 2010s, for better or worse, vast conglomerates like Facebook and Google effectively take on the role of global communication regulators. Mark Zuckerberg’s repeated insistence that Facebook is nothing more than a technology company is a sad reprise of Fowler’s laissez-faire ideology. It is also deeply hypocritical, for the firm imposes all manner of rules and regulations on both users and advertisers when those norms generate profits for it (Pasquale, 2015b).

The public sphere cannot be automated like an assembly line churning out toasters. As Will Oremus has explained, there are aspects of the journalistic endeavor that are inherently human; so, too, are editorial functions necessarily reflective of human values (Oremus, 2014). To be sure, there will be deep and serious conflicts over the proper balance between commercial interests and the public interest in assigning prominence to different sources; in deciding how much transparency to give decisions made about such issues; and how much control individual users should have over their newsfeeds, and the granularity of that control. But these are matters of utmost importance to the future of democracy. They can no longer be swept under the rug by plutocrats more interested in stock returns and artificial intelligence advances than the basic democratic institutions and civil society that underpin each.

References


Citron, D. *Hate crimes in cyberspace*. Cambridge, MA: Harvard University Press.


Lubbers, E. (2016). There is no such thing as the Denver Guardian. *Denver Post*, November 5. denverpost.com/2016/11/05/there-is-no-such-thing-as-the-denver-guardian/


Pasquale, F. (2010). Beyond competition and innovation: The need for qualified transparency


Pasquale, F (2015a). Digital star chamber: Algorithms are production profiles of you. What do they say? You probably don't have the right to know. Aeon, August 18. aeon.co/essays/judge-jury-and-executioner-the-unaccountable-algorithm


net: Race, sex, class and culture online (pp. 147-159). New York: Peter Lang.


Stucke, M. E. & Grunes, A. P. (2017). Data-opo-
papers.cfm?abstract_id=2927018


Suskind, R. (2004). Faith, certainty and the pres-
idency of George W. Bush. New York Times, Oc-
tober 17. nytimes.com/2004/10/17/magazine/
faith-certainty-and-the-presidency-of-george-
w-bush.html

Tufekci, Z. (2016). Mark Zuckerberg is in de-
com/2016/11/15/opinion/mark-zucker-
berg-is-in-denial.html?_r=2

book Op-Ed is a joke, Gizmodo, November 30.
gizmodo.com/new-york-times-public-editor-
slams-facebook-op-ed-1789511334;

Winston, J. How the Trump campaign built an identity database and used Facebook ads to win the election. Medium, November 18. medium.com/startup-grind/how-the-trump-
campaign-built-an-identity-database-and-
used-facebook-ads-to-win-the-election-4ff-
7d24269ac#.4oaz94q5a

Winters, J. (2011). Oligarchy. New York: Cam-
bridge University Press.

Related Publications

PLATFORM COOPERATIVISM
Challenging the Corporate Sharing Economy
Trebor Scholz - January 2016

www.rosalux-nyc.org