

A Shattered Public Sphere

HOW SOCIAL MEDIA PLATFORMS TRANSFORM THE POLITICAL DEBATE

Mirko Tobias Schäfer

Introduction

Once, social media were heralded as means for emancipating citizens, for balancing power asymmetries and allowing common people to inform themselves and express their opinions. Today, we associate social media with fake news, hate speech and manipulative algorithms. After more than two decades of the World Wide Web, social media have emerged as powerful brokers of attention, channelling access to information and connecting audiences. The platform providers appear in parliamentary hearings as deceitful as they seem to be reluctant to fix their problematic services. After Brexit, Trump and the emergence of right-wing populism, policy makers and commentators argued that social media platforms undermine the open society through constituting filter bubbles and disseminating fake news and hate speech. This chapter explains how we went from being enthusiastic about social media to being disappointed. But it is not enough to blame social media platforms for the current situation. We have to look at the interplay of technological affordances and media practices, economic interests and institutional change. Looking at the notion of the public sphere and how it shaped our understanding of engaging in political debate and participating in an open society, this article emphasizes our far-fetched expectations of technology as an engine of democratic progress. Introducing the term of implicit participation, I show how social media platforms were successful in implementing media practices into easy-to-use interfaces and in channelling user activities. It cannot be emphasized enough that this constituted the engagement of large audiences indifferent to the inherent values of the public sphere and the open society. Rejecting the notion of the filter bubble, I show how users engage with media content and how mainstream media are engaging with social media. In conclusion, this chapter exposes the policy reactions to the perceived threats of fake news and hate speech as inappropriate, inefficient and actually damaging for the open society and the public sphere.

Participating in the Public Sphere

The overly optimistic perception of social media was very much informed by our understanding of the public sphere and the values of citizen participation in political debates. Referring to the early 19th century, German philosopher Jürgen Habermas described how common citizens debated public affairs in cafés, reading clubs and associations, how they used media such as leaflets, newspapers and letters to share information. He defines this public sphere as a "network for communicating information and points of view" (quoted in Castells 2013). This public sphere is an essential part of the open society, where freedom of speech shapes opinion forming and affects democratic deliberation. However,

throughout the 20th century, this public sphere appeared to be dominated by mass media: newspapers, film, radio and television. Mass media have been seen not only as gate keepers mediating awareness of issues and disseminating information to audiences, but also as monopolistic agents of power which could as well manipulate public opinion forming. Hence, the emergence of computer technology and the World Wide Web has been lauded as an opportunity for common citizens to balance the power of mass media and foster the open society. The Internet and the World Wide Web have been seen as means for connecting people around the globe, disseminating information and facilitating a low threshold access to information and participation in media production, and political debate and opinion-forming (Schäfer 2011). The perception of social media as means for balancing power asymmetries in media society was very much informed through our understanding of the public sphere and the open society. The emergence of the World Wide Web has been welcomed as such emancipation and manifested even in the popular Time Magazine cover displaying “You” as person of the year 2007. The tagline read. “Yes, you. You control the information age. Welcome to your world.” This notion of participation was not new at all.

Throughout the 20th century, political activists, artists, and scholars considered the massive participation of citizens in media production and political debate as essential for democracy to thrive. In 1926, playwright, director and theater theorist Bertolt Brecht criticized the state regulation of radio sets and airwaves as constituting an anti-democratic mono-directional apparatus. As radio inherently allows for two-way communication, Brecht suggested to deliberately use radio sets as communication and not as broadcasting devices. He dreamed of a medium for connecting people to debate public affairs (Brecht 1999). In his own theatre practice, Brecht developed a concept that would involve the audience, turning them effectively from onlookers into participants (Benjamin 2003).

This plea for participation informed many media theories of the 20th century. Enabling audiences to participate has been understood as emancipation of passive onlookers, balancing power asymmetries in media production and was generally understood as beneficial to political debate and democracy (Schäfer 2011). Unsurprisingly, the new media were seen as the long-awaited facilitation of these requests. They were seen as means to enable common citizens to inform themselves despite social, economic or geographic disadvantages. Many Silicon Valley companies echoed this promise: Facebook’s mission statement is to connect people and bring the world closer together, allow users to share and express what matters to them.¹ Similar could be said for the networking company Cisco Systems (see Schäfer 2011:31-34) and other companies. We attributed qualities to the new technologies that would connect with our own values of engaging in the public sphere and deliberation in an open society. In order to understand the misperception of the new technologies as inevitable means for democratization and

¹ Facebook mission statement: < <https://newsroom.fb.com/company-info/>>

social progress we need to look at the common understanding of participation.

Explicit participation and implicit participation

Habermas and Brecht understood participation as a conscious and explicit activity, where informed citizens would engage in matters of public interest.² Key to Habermas' notion of the public sphere was the rational citizen, whose political opinion is shaped through reasonable debate, arguments supported by factual information and media that served not only as gate keepers for mediating topics of political discussion, but also as channels for receiving verifiable information. Another element essential to this public sphere is the mutual respect of the participants who despite their different political positions still respected their opponents as fellow citizens, and trusted that reason and factually correct information forms the basis for political opinion forming. We have understood the World Wide Web and its many applications as a mere extension of the already existing mass media, newspapers, radio and television, though with a lower threshold to participate. Before the advent of affordable personal computers, internet connections and software for editing texts, images and videos, producing media content required skills and training, and was costly in reproduction and dissemination. If citizens engaged in these activities it was obviously an explicit participation in media production and the public sphere. While it was relatively easy to build websites, send emails, use mailing-lists or web forums and other popular applications of the World Wide Web during the late 1990's and the early 2000's, it became even more accessible with social media platforms. And it is here, where next to explicit participation a form of implicit participation emerged (Schäfer 2011:46). In contrast to explicit participation, implicit participation describes how user activities are channelled through graphical user interfaces and monitored through real-time analysis of interaction data.

Social media platforms effectively implement many media-practices that were developed during the first ten years of the World Wide Web into easy-to-use interfaces: editing websites, videos, images, connecting with other users, and sharing content became features that did not require any knowledge of html, file servers, or software programmes. These platforms enable the participation of users who were widely computer illiterate as the interfaces made using the technology even easier than it already was (Schäfer 2011). This low threshold is relevant for understanding the success of social media platforms. Not only are their services usually free, they also persuasively stimulate use of the platform and new connections with other users. This manifests most clearly in buttons for disseminating or liking content: the retweet button on Twitter allows users to disseminate content created by others. The same feature is visible on

² It goes without saying that Brecht and Habermas, as well as many other philosophers, were profoundly informed by the European tradition of Enlightenment, the "emergence from self-imposed immaturity" as Immanuel Kant formulated it in 1784.

Pinterest (repin), Instagram, Facebook, LinkedIn YouTube (share) and other platforms. Buttons displaying stars, hearts or thumbs-up are placed to allow users to easily indicate appreciation for content they find online, although the favourite button (now a heart) in Twitter is also used as bookmark. It cannot be emphasized enough that these buttons facilitate a low-threshold for dissemination and interaction of users. They allow people to make use of the platform without actually creating content themselves. As most users probably have difficulty making use of only 140 characters to create a meaningful message, it is utterly convenient for them to be able to just disseminate what others have created. Platform providers utilize data gathered from user activities and their connections to organize and disseminate content on their platforms.³ Simply through using the platform, users provide valuable information about their interests, social connections, and their habits. Interaction data inform platforms in real-time whether content is sparking high engagement and within which clusters of users the content is spreading. Analysing and using these data for targeted advertising, facilitating channelled dissemination of content and brokering access to audiences has become a lucrative business model of platform providers. The data are the business model, and the algorithms are attempts to use them in ways which serve the business model best.

Revisiting Filter Bubbles

The term filter bubble was coined in a TED talk by Pariser (2011). Describing his personal experience, he argued that browser history affects the information users receive through Google, Facebook, YouTube and other platforms. As a result, users would only receive information fit to their individual profiles and effectively shield them from other perspectives. Eventually it would create a situation where users perceive the world purely through this ‘filter’. Despite lacking empirical evidence, filter bubbles are now perceived as a threat to a pluriform media landscape. However, research indicates that there is no such thing as a filter bubble (Borgesius et al 2016).⁴ Our research at Utrecht Data School and Datafied Society confirms this (Wieringa et al 2018). Looking at a two-week sample of Dutch Twitter, we can confirm the existence of a large number of topic communities. But topic communities are not filter bubbles. Within their tweets users refer to a variety of other platforms, e.g. YouTube, Facebook, newspapers, blogs, or other sites. When looking at political topic communities, e.g. right-wing oriented users, we noticed

³ If you were wondering why you are often asked to identify cars, storefronts, traffic lights, pedestrians or road signs when you are about to enter a website, it is because you are implicitly participating in training Google’s AI for self-driving cars.

⁴ See also a report commissioned by the Commissariaat voor de Media: Moeller, Judith, Natali Helberger and Mykola Makhortykh: Filter Bubbles in the Netherlands. Commissariaat voor de Media, 2019 <<https://www.cvdn.nl/nieuws/geen-indicatie-voor-bestaan-van-filterbubbles-in-nederland/>>

that they do not exclusively refer to media promoting their own ideological perspectives. Among the right-wing cluster in our sample, we found that left-leaning Volkskrant was ranked among the top four referenced URLs. The first three spots were held by Telegraaf, followed by The Post Online and GeenStijl. GeenStijl is a populist weblog and The Post Online a clearly right-leaning online news format. Looking at the topic community we characterized as politically centre-left, we found the top four referenced URLs were the public media service site NOS.nl, the left-leaning Volkskrant, YouTube, and Telegraaf, ranked four. Within the overall sample, YouTube and Facebook are most often referenced in tweets, followed by the tabloid the Telegraaf, the Belgium news page Het Laatste Nieuws, Sport in Nederland, the Dutch news programme NOS, NPO Radio, the news site Nu.nl, the tabloid Algemeen Dagblad, and in tenth place, Google.com. This image does not show political preferences as much as it shows a preference for social media and tabloid formats (see figure 1).

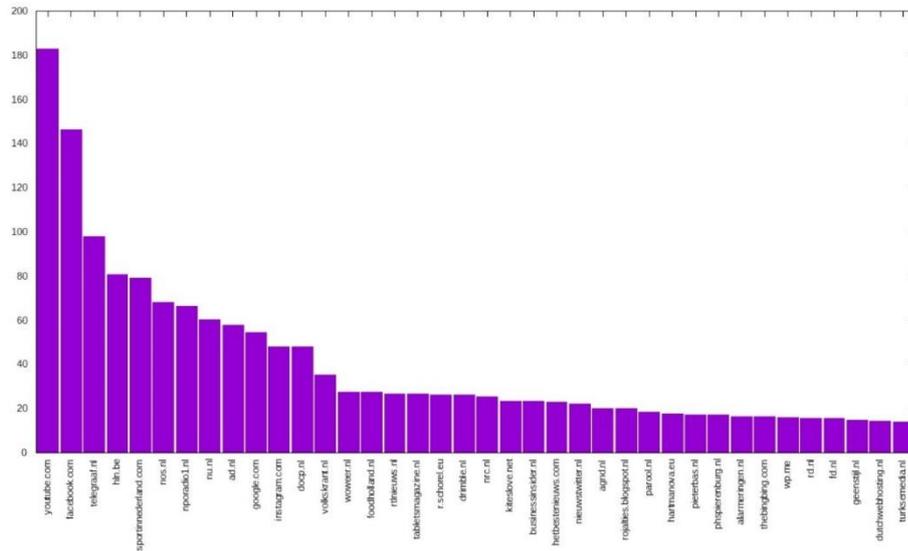


Fig 1: Most frequently shared URLs; total number of referenced domains: 58,662; total number of users: 461,279.

Even within ideologically connotated topic communities, users refer to links to media that expose political positions that are contrary to their own network. What is more apparent is that across the entire sample the most frequently shared URLs seem to be tabloid media. Maybe we should perceive social media as being more compatible with messages that traditionally run in tabloids. Research indicates that social media platforms seem to favour sensational and emotional messages. As the algorithms for dissemination are intrinsically informed by user activity, they naturally prefer content that constitutes engagement of users (likes, retweets, comments, ...). The low threshold to participate in the dissemination of messages also connects well to instant reactions. Users

who are emotionalized, shocked or excited about messages they receive can share them in an instant, maybe in a moment when they are more stimulated by emotional instincts than reason. If that were true, an essential aspect of Habermas' public sphere would be bypassed. Communication about political affairs was to be informed by facts, reasonable argument and mutual respect of the participants. These elements can go missing completely in social media conversations, where messages are distributed that defy reason as much as factual soundness, and where the tone of conversations often neglect manners or respect. Policy makers emphasize that there is a corporate responsibility to prevent misinformation and hate speech being disseminated through these platforms. However, pointing solely to platform providers as culprits is not sufficient.

It's complicated: Mainstream media and social media

Channelling news and brokering awareness of large audiences, social media platforms have garnered a powerful position within the media landscape. A large percentage of the younger population consumes news solely through social media platforms⁵. Unsurprisingly, traditional media perceive these platforms as competitors who steal both their content and their audiences. However, there is a dynamic at play which connects the web platforms with traditional media. It manifests in the ways journalists and editors in chief at traditional media outlets use and interact with social media platforms. Populists such as Trump, Wilders, Salvini, Baudet and others are heralded as extremely social media savvy. They do not only have a large following on their social media accounts, they also send messages that are eagerly disseminated by their followers, but unfortunately also by journalists. By doing so, mainstream media amplify conversations from social media. Newspapers and broadcasting media also devote too much attention to social media which they seem to perceive as windows to current affairs or sentiment. It manifests in their practice to use social media messages as vox populi in their own reporting, but also in the choice of topics they cover to the extent that the process of gate keeping -selecting which issues to cover and which news to report- seems to be delegated to social media. This is problematic as social media are not representative and display a bias in their conversations. Emotionalizing issues, sensationalist content, and polarising positions are overly represented in social media while factually informed debate, nuanced positions and reason are marginalized. How well traditional media and social media connect became visible in an incident in Austria, where in 2018, Johann Gudenus, deputy leader of the far-right Freedom Party, accused a young refugee of sympathizing with a terror group. That unfounded accusation was then reported by the tabloid Kronen Zeitung without prior fact-checking. Gudenus then shared this fabricated 'news' on

⁵ According to this survey by the American Press Institute:
<https://www.americanpressinstitute.org/publications/reports/survey-research/millennials-news/>

social media again.⁶ It also indicates that misinformation is not exclusive to social media. It can be traditionally found in tabloid formats. Looking at a collection of covers of the US American tabloid *National Enquirer* during the presidential election campaign reflects exactly the same sentiment and the conspiracy theories about Hillary Clinton circulating in social media and the US right-wing blogosphere. It appeared that the publisher deliberately pushed the Trump platform, which debunks the argument that Trump came to power thriving on social media use alone.⁷ Traditional media, especially the *National Enquirer* and Fox News were instrumental in amplifying Trump's rhetoric.

Sometimes it even goes further. Notorious 'hate queen' Katie Hopkins commented on the shooting incident that took place in March 2019 in Utrecht. The video message she recorded on her phone for her large social media following was broadcast that very evening on one of the Netherlands' most influential talk shows, *Pauw en Jinek*.⁸ The talk-show hosts introduced her as "very conservative journalist" despite a lack of journalistic training or established media affiliation. Her racist message directed to her own audience was presented as yet another but "very conservative" [sic!] point of view on the Utrecht events. In May 2019, far-right leader Thierry Baudet shared a video from the neo-Nazi group *Die Identitären* which argued that immigration will lead to massive rape of women in Europe, and accused political leaders of having neglected this in the same way Germans had been ignorant of the fate of Jews during the Nazi regime.⁹ Without pointing out the origin of this video, it was broadcast prominently on public television during a TV debate between prime minister Mark Rutte and Party Leader of the Forum for Democracy Thierry Baudet.

Mainstream media amplify social media. But the relations go further. Out of a flawed understanding of fair and balanced reporting, qualitative journalism seems to include extreme perspectives within their coverage, even if they represent only marginal positions and completely unfounded claims. Unfortunately, journalists often fail to contextualise appropriately, and question the unfounded claims. We also see that tabloid media connect well with social media; tabloid content is shared more often and

⁶Die Presse: FPÖ-Vorwürfe gegen Lehrling falsch: Keine Entschuldigung von Gudenus, 6.9.2018, online:

<https://diepresse.com/home/innenpolitik/5491928/FPOeVorwuerfe-gegen-Lehrling-falsch_Keine-Entschuldigung-von-Gudenus>.

⁷ <https://www.politico.com/magazine/story/2017/01/tabloid-newspapers-trump-media-propaganda-214627>

⁸ Pauw en Jinek: De verkiezingen, 18.3.2019, online <https://www.npostart.nl/pauw-jinek-de-verkiezingen/18-03-2019/KN_1705873>, from 39:33 onwards.

⁹Burger, Peter: Baudet promootte video van extreem-rechtse antisemitische site uit Nederland, Nieuwscheckers, 22.5.2019, online: <<https://nieuwscheckers.nl/nieuwscheckers/baudet-promootte-video-van-extreem-rechtse-antisemitische-site-uit-nederland/>>.

their topics (conspiracy theories, climate denial, claims about mass immigration) appeal to social media audiences.¹⁰

Wrong and dangerous policy decisions

Brexit, Trump, the emergence of populist groups and parties across Europe challenge our democratic system. Politicians and traditional media consider social media platforms as essential part of the problem. Populist politicians, such as Trump, Salvini, Wilders, Baudet and others seem to thrive on social media. Their activities and their audiences are widely associated with hate speech and fake news. Meanwhile, policy makers push for far-reaching solutions to ban hate speech and fake news. In Germany, the Network Enforcement Act was introduced in 2018. The law demands platform providers to remove or block access to content which obviously violates German law within 24 hours after notification. In more complex cases the time limit is seven days. The proposal was heavily criticized because it would force providers to remove more content than necessary. Most providers have already applied their own guidelines for content, which also are not necessarily in line with what is allowed under free speech laws.¹¹ However, politicians are still tempted to tame the glut of messages. In many European countries, various bills are under way to stifle the dissemination of hate speech. Proposals range from a requirement that users post only under their real names, to the use of filters to identify illegal content. Most recently, the European Parliament voted for a new Copyright Directive. The highly contested proposal was heavily lobbied for by copyright holders, mostly publishers and media corporations. The law is supposed to prevent social media platforms from disseminating content without compensating corporate copyright holders. They are also required to control which content their users are uploading. Although the law deliberately avoids the term uploadfilter, automatized monitoring of user uploads is inevitable in order to comply with the law.¹² There are already automatized filters at work on these platforms. YouTube's ContentID, which cost 60 million US dollars to develop, compares uploads with a database of commercial content from the music and film industry. But they routinely block access to perfectly legal content because they are unable to distinguish between satire, persiflage, and the many exemptions from copyright allowed under fair use. Critics

¹⁰ Tabloid media also adopt distinct formats from the blogosphere. In the Netherlands; the populist weblog GeenStijl is owned by the Telegraaf, and The Post Online was owned by Veronica Ventures. Both sites are predominantly present in conversations within right-wing politics topic communities.

¹¹ After a year in effect, it appears that only a minority of the reported content was actually illegal. The complaint centre for implementing the Network Enforcement Act received 8617 complaints in 2018, and less than 50% of these were related to illegal content. Zeit online:

Netzwerkdurchsetzungsgesetz findet kaum Anwendung, 12 March 2019: <<https://www.zeit.de/politik/deutschland/2019-03/netzdg-netzwerkdurchsetzungsgesetz-jahresbericht-eco-beschwerdestelle>>.

¹² Every minute 300 hours of content are uploaded to YouTube. Every minute on Facebook, users share 317,000 status updates; 147,000 photos and 54,000 links. <<https://www.omnicoreagency.com/facebook-statistics/>>.

of the law included the chief data protection office of Germany, the UN rapporteur on freedom of speech, associations of scientists, the inventor of the World Wide Web and almost every expert on the topic. However, their criticism was not well covered by mainstream media, and their arguments were not represented in a balanced way in newspapers and broadcasting services. In March 2019, as the final vote in the European parliament neared, critics mobilised hundreds of thousands of people to take to the streets, where they also collected the biggest number of signatures on a petition for the European Parliament. The copyright directive is an important example, as it indicates policy makers' trust in automatized monitoring of content and cooperation with large platforms. Although the entire lobby campaign for the copyright directive was framed as claiming compensation from Google and Facebook for creatives, it actually catered to the interest of those two companies. Both Facebook and Google benefit largely from the new law. They have the power to negotiate directly with large copyright holders, and they have the money to develop algorithmic solutions to monitor and filter user content. Most importantly, the law is a burden for small and medium companies, and will keep competition against the dominant platforms at bay. They can license their filters to other platform providers who do not have the means to develop their own solution. What Facebook, Google and Amazon do not have is a commitment to the open society, democracy and free speech. Their services function in a dictatorship just as well as in a democracy. They have no obligation or interest to facilitate a public sphere. Without justification they can delete content, exclude users, and channel awareness as they see fit. But they are considered important partners not only in the fight against copyright violations; the EU's illegal content initiative is also promoting actions against the dissemination of terrorist content. The proposal for the controversial "EU regulation on terrorist content online" requires platform providers to take down identified terrorist content within one hour.¹³ While this seems to be an effort that is even challenging for large and well-funded corporations such as Facebook, YouTube or Amazon, it is absolutely not feasible for non-commercial platforms. A preview on the reality of the EU's terrorist content policy was given recently by the French Internet Referral Unit. Tasked with identifying malicious content online, they falsely flagged 550 URLs of the Internet Archive as terrorist content.¹⁴ The Internet Archive is one of the biggest non-profit platforms for archiving public domain content. Among the falsely flagged items were the prestigious Prelinger Archive, a large collection of films related to US American culture, the song archive of the Grateful Dead, and the Gutenberg Project, a large collection of public domain literature. If the one-hour deadline would have been in effect already, the understaffed Internet Archive would have

¹³ Jon Porter: Upload filters and one-hour takedowns: the EU's latest fight against terrorism online, explained, The Verge, 21.3.2019, online: <<https://www.theverge.com/2019/3/21/18274201/european-terrorist-content-regulation-extremist-terreg-upload-filter-one-hour-takedown-eu>>.

¹⁴ Chris Butler: Official EU Agencies Falsely Report More Than 550 Archive.org URLs as Terrorist Content, Internet Archive, online: <<https://blog.archive.org/2019/04/10/official-eu-agencies-falsely-report-more-than-550-archive-org-urls-as-terrorist-content/>>.

been forced to just take down the flagged URLs and would have effectively cut off users from accessing perfectly legal content. In this case, the problem is not only the blatant incompetence of the French authorities, but the policy makers' ignorance of media practices and technology.

Efforts to tame hate speech and combat fake news or terrorist content online are poorly informed in terms of technological feasibility and how users actually use the internet. Most politicians seem unaware that the internet does not only consist of Facebook and YouTube, but a multitude of non-profit platforms, small and medium-sized commercial providers, and millions of users participating daily in the creation of content and the dissemination of messages. The policies underway threaten to stifle opportunities for common users to exchange information and points of view online. The simplistic view many policy makers have of the internet and how it is used does not reflect its actual nature.

Bibliography

- Benjamin, Walter. Der Autor als Produzent. In Walter Benjamin, *Medienästhetische Schriften*, pp. 231-247. Frankfurt a.M.: Fischer, 2003 [1934].
- Borgesius, Frederik, et al. "Should we worry about filter bubbles?" *Internet Policy Review. Journal on Internet Regulation* 5.1 (2016).
- Brecht, Bertolt. Der Rundfunk als Kommunikationsapparat. In Lorenz Engell et al. (eds.) *Medienkultur. Die maßgelichen Theorien von Brecht bis Baudrillard*, pp. 259-263. Stuttgart: DVA, 1999 [1932].
- Castells, Manuel. *Communication Power*. Oxford: Oxford University Press, 2013.
- Habermas, Jürgen. *Strukturwandel der Öffentlichkeit*, Frankfurt a.M.: Suhrkamp, 1962.
- Habermas, Jürgen. *The structural transformation of the public sphere: An inquiry into a category of bourgeois society*. Cambridge, MA.: MIT Press, 1991.
- Pariser, Eli. *The filter bubble: What the Internet is hiding from you*. Penguin UK, 2011.
- Schäfer, Mirko Tobias. *Bastard culture!: how user participation transforms cultural production*. Vol. 6. Amsterdam University Press, 2011.
- Wieringa, Maranke, Daniela van Geenen, Mirko Tobias Schäfer, and Ludo Gorzeman. 2018. *Political topic-communities and their framing practices in the Dutch Twittersphere*. In: *Internet Policy Review*, Vol 7, No 2.