Who moderates my social media? Locating Indian workers in the global content moderation practices
Ahmad, Sana

```
●●●●●●   Digital
●●●●●●●  Communication
●●●●●●●
●●●●●●   Research.de
```

Abstract: Building on the growing concerns around hate speech and harmful content on social media, this chapter analyzes the processes by which content is moderated on leading social media platforms. The outsourcing practices of platform operators or social media companies to acquire content moderation services from third-party companies have been acknowledged in the public discourse. Details regarding these outsourcing relationships and power mechanisms remain obfuscated, however. Using empirical data from India, this chapter presents a global value chain perspective on the mechanisms by which US-based social media monopolies source content moderation services from Indian information technology business process outsourcing (IT BPO) supplier companies. The agreements established between the two parties direct the content moderation labor process through which Indian workers' labor power is transformed into productive labor.

Sana Ahmad

Who Moderates My Social Media?

Locating Indian workers in the global content moderation practices

1   Fixing "our broken" social media

In a recent media article in *The Guardian*, technology reporter Julia C. Wong put together a list of proposals by North American researchers and activists to fix "our broken" social media (Wong, 2021). These proposals rather being speculative underscore concrete actions to regulate social media platforms. "We cannot fix what we do not understand," notes one of the experts, Alex Abdo, litigation director at the *Knight First Amendment Institute* (Wong, 2021). Abdo advocates for enabling independent inquiry by researchers and journalists to explain how social media companies have managed to prioritize user retention at the cost of allowing hate speech and fake news to circulate on their platforms.

The calls for social media companies to have transparent content moderation policies and practices on their platforms have gradually increased over the last several years. Several nation-state governments today, especially with evidence of the use of social media to influence their election results, are using legal routes to prohibit the presence of hate speech, fake news, and other propaganda on social media. Furthermore, countries such as the United Kingdom (UK), India, and

others are in the process of delegating the responsibility of monitoring and controlling these spaces to social media companies. However, putting the onus of social media management on companies and privatizing law enforcement can have repercussions for users' freedom of speech, as has been pointed out by many, including Brigitte Zypries, in her former role as Minister of Justice in Germany (Agence France-Presse, 2017).

The positions offered in this chapter do not argue for or against regulation of speech on social media platforms. Instead, I take Abdo's proposal seriously on understanding social media and platform operations before trying to fix it. The material conditions underlying the functioning of social media platforms and the built-in power asymmetries are the focal points of this chapter. Drawing heavily from the labor process debate, the commercialized practice of content moderation is examined here, with specific attention placed on the working conditions of content moderators employed at third-party contracting companies in India. Noting the contemporary public discussions and significance assigned to the function of content moderation for social media, this chapter aims to motivate the reader to consider the ongoing treatment of content moderation practices as *industrial secrets* by social media companies.

However, this chapter does not chart a distinct relationship between content moderation production processes and the proliferation of hate speech on social media. Increased public attention to harmful content on these platforms has elicited, on the one hand, techno-solution-oriented responses and, on the other, the assurance of contracting additional human reviewers by social media companies. While transparency reports from global social media monopolies show that the prevalence of hate speech has reduced on their platforms, it cannot be confirmed whether this has been made possible using content moderators' labor power or through the exclusive application of automated filters and technologies[1].

---

1   The fourth quarterly reports from 2020 can be accessed on the official websites of Google and Facebook. The 'YouTube Community Guidelines Enforcement' report is accessible at https://transparencyreport.google.com/youtube-policy/removals, and Facebook's 'Community Standards Enforcement Report' is accessible at https://transparency.facebook.com/community-standards-enforcement. The biannual 'Twitter Transparency Report' is accessible at https://transparency.twitter.com/en/reports.html.

The analysis presented here is derived from research fieldwork in India, which I undertook as part of my doctoral inquiry. In total, 35 guided interviews were conducted with target participants in India. The chapter consists of the following sections. I start by defining content moderation and the commercialization of this practice. Following this, I attend to the question of why it is obfuscated from the public view. I then present an overview of the content moderation labor process and possibilities for resistance, if any. Finally, I conclude by underlining the importance of further research and the relevance of policies in regulating these outsourced practices.

2   Content moderation: Why it matters

The practice of content moderation follows a pattern of evolution similar to that of the internet-based services. With an increase in the commercial application of the Internet in the 1990s, and an expansion of Internet-based services, the need to screen and monitor these services grew as well. Commercial services based on the World Wide Web, such as email services (Hotmail.com, Yahoo, AOL, etc.), classified advertisement services (Craigslist), dating services (Match.com), and peer-to-peer file sharing services, were monitored and controlled according to local regulations and company standards. Information scientists and inter-personal communication researchers were quick to identify the growth of social media as a "computer-mediated communication" in the form of emails, forums, and Bulletin Board Systems (Rice, 1980; Kerr & Hiltz, 1982 in Burgess et al., 2018). Many of these text-based social communities, followed by an increasing number of social technologies in the 1990s and mid-2000s (MySpace, Wikipedia, Reddit, etc.), placed emphasis on online community management through open and voluntary moderation (Roberts, 2017).

The shift of focus on social media from *social network sites* (boyd & Ellison, 2007) to *social media platforms* (Gillespie, 2018b) accompanied a surge across several disciplines, including media and communication studies, to examine the ethics of data culture, especially the collection, monitoring, and monetizing of user-generated data by social media companies (Herman, 2014; Helmond, 2015). It was Roberts (2019), however, who, through her empirical investigation, was able to link large-scale social media platforms in the United States of America (USA)

with the commercialized practice of content moderation. According to her definition, "commercial content moderation is the organized practice of screening user-generated content posted to Internet sites, social media and other online outlets, to determine the appropriateness of the content for a given site, locality, or jurisdiction" (2017, p. 1).

In a similar vein, Gillespie (2018a) identified synchronously occurring processes of content moderation on social media platforms and public exchange. Similar to Roberts (2019), Gillespie considered content moderation as the core process for maintaining these platforms, and he goes on to equate it as an "essential, constitutional and definitional" function of social media platforms (Gillespie, 2018a, p. 21). However, much before content moderation as a commercial practice could receive scholarly attention, investigative articles in the media exposed its outsourcing to peripheral states in the USA and later its offshoring to geographically dispersed locations across the world (Stone, 2010; Chen, 2012; Chaudhuri et al., 2014). India, along with the Philippines, has been observed as crucial locations for content moderation outsourcing.

3 A closer look at the hidden practices of content moderation

In his seminal work on providing a historical materialist understanding of *digital materialism*, Gottlieb examined the "mystification or metaphysical obfuscation" of processes associated with digital technologies (Gottlieb & Karatzogianni, 2018, p. 2). Gottlieb's focus on digital materiality allows us to acknowledge the often-times hidden labor that goes into producing and maintaining the technologies of today. Further, it prompts us to investigate the underlying social relations that constitute the technological processes. Examining the political economy of digital media certainly opens new opportunities for studying the unpaid activities of social media users and their commodification by social media companies (Fuchs, 2014, 2010; Dyer-Witheford, 2010). However, Gandini argued that such a broad analysis of labor and digital technology could risk understudying the hidden dimensions of *digital labor* (2020).

Gandini proposed considering platforms, including social media, as organizational actors and examining the "manifold ways in which the capital-labor relationship is enforced through them" (2020, p. 9). Correspondingly, the focus of this

essay is on explaining the production model of content moderation, which remains, as we have determined before, an essential feature of social media platforms. Much of this content moderation work, which involves screening large amounts of user-generated information within a very short amount of time, is carried out by outsourced workers located in the world peripheries.

Content moderation practices and the outsourcing of labor are hidden from the public view, thus making it difficult for independent researchers and journalists to assess these processes. Some of the self-described motivations of social media companies to maintain this secrecy are as follows: to protect the identities of workers (Gillespie, 2018b), to prevent the users who post illicit content on social media platforms to "game the rules" (Roberts, 2016, p. 7), and to "guard the proprietary tech property and gaining cover from liability" (Buni & Chemaly, 2016, p. 12).

In the wake of leaks in media articles as well as lawsuits filed by content moderators against social media companies, the industry's secrets are spilling out. Yet, public and legal focus on the hidden labor of content moderation has remained rather limited. As mentioned above, national legislation in different countries is taking shape in trying to shift the liability on social media companies for hosting illegal online content, disputing the protection of these companies under the *safe-harbor* legislation in the USA.[2] While these developments have certainly allowed us to challenge what some have called a "marketplace orientation" of Section 230 (Medeiros, 2017, p. 2), they have yet to take into concern the production process of content moderation and labor, which goes into sustaining this essential practice.[3]

The outsourcing of content moderation work by social media companies has created global content moderation value chains. While the content moderation policies and software are designed within social media companies, the actual labor of content moderation, which is often low-paid and "rote, repetitive, quota-driven, queue based" (Roberts, 2019, p. 92), is outsourced to contracted content moderators who are placed at great distances from these companies. In public discourse, content moderation has often been understood as an automated task, and the reality of human content moderation has only been explored in the

---

2   Section 230 of the Communications Decency Act in the United States provides the Silicon Valley-based social media giants, along with other websites, a safe harbor from liability for user-generated content or third-party content posted on their platforms.

3   Medeiros notes that for these companies, "suppression of speech can be anathema to the marketplace theory" (2017, p. 2).

recent years. Studying the offshore practices of content moderation on social media platforms is challenging. Most notably, the term "content moderation" is not a standard business terminology. Instead, several other job titles, such as "system analyst," "website administrator," "process executive," and others, are assigned to moderators by supplier companies in India (Ahmad & Krzywdzinski, 2022).[4] Roberts (2019, p. 40) noted that these "multitudinous" job titles function to further conceal the content moderation process.

The deliberate concealment of this process by target social media companies located in the Global North and complying supplier companies in India compels me to argue that the rules governing outsourcing relationships and the resulting labor processes of content moderators in India are designed to create opacity around content moderation practices. As we will see in the following section, social media companies outsource content moderation to India (our target location) through traditional business process outsourcing practices in which gig work online platforms do not play a major role. Most content moderation labor processes are organized and controlled by social media companies, their standards, and software infrastructures.

4   Exploring the labor process of content moderation

Over one-tenth of moderation workers worldwide are located in India, which is one of the main destinations of content moderation outsourcing.[5] In my research, I identify content moderation as a back-end, non-voice business process that is supplied as a service to their clients, including to social media companies, by information technology business process outsourcing (IT BPO) sector companies in India.

---

4   The suppliers referred to here are information technology business process outsourcing (IT BPO) companies who provide a range of services and technological solution to their clients located around the globe. A motley assortment of clients requires content moderation services for their social media platforms, e-commerce websites or simply their user-content hosting websites. The Indian IT BPO companies supply content moderation services to these different clients.

5   The estimation was made by Himanshu Nigam, former chief security officer at the social media platform MySpace and a former security executive at Microsoft (Chaudhuri et al., 2014). There are no publicly available statistics which could show the impact of content moderation services on the Indian labour market.

The analysis presented in this chapter is based on research fieldwork in India, and grounded theory methodology guided both the data collection and data analysis processes. The research was undertaken between January 2019 and April 2019, and constituted nine interviews with content moderators and three interviews with content operators[6], six interviews with management at the supplier companies in India, two interviews with domestic social media companies, seven interviews with trade unions, and eight interviews with civil society organizations. Moreover, informal meetings were held with experts in the fields of labor law, technology, and free speech to achieve more insights into this service work.

Gaining access to the participants was extremely challenging and explained the absence of representation from international social media companies. Most of the workers were approached on an international networking website for professionals, and the rest were contacted using the snowball sampling technique. Considering the sensitivity of the subjects, great care was taken in protecting the identities of all participants, both during and after the data collection process.

With content moderation practices treated as industrial secrets, as has been described before, the brief description of content moderation outsourcing mechanisms presented here is influenced by the vast body of literature on Indian call center companies (within the IT BPO sector) and the work organization and management strategies of these companies. The discussion on labor processes in call center companies highlights the subordinate position of Indian companies in global value chains and the subsequent vulnerability of the workforce (Batt et al., 2005).

Global content moderation value chains are facilitated by service level agreements (SLAs), which are established in this case between social media companies in the Global North and content moderation suppliers in India. Depending on the terms of the agreement, specific tasks are allocated to social media companies and their suppliers. Training, developing content moderation policies, and other product-oriented tasks are managed by social media companies.[7] By contrast, tasks such as managing wages, leave of absence, workplace

---

6   *Content operator* is an official designation at domestic and regional social media firms, wherein the workers are assigned content-related tasks, such as user acquisition and retention, along with either moderating the content themselves or overseeing the moderation tasks done by external freelance moderators.

7   Product-oriented task refers to social media platform as a product which is designed by its proprietor, the respective social media company.

conflict, and other human-resource related tasks are handled by the supplier companies. These agreements are mostly project based and are determined by measurable standards, such as quantity targets (the amount of user content moderated) and time. Such factors enable flexibility and scaling-up opportunities (regarding the volume of their outsourced content moderation business) for social media companies.

As observed, these content moderation value chains are characterized by high power asymmetries between social media companies on the one hand and supplier companies on the other.[8] These, I argue, have an influence on the content moderation labor process. The different kinds of content moderation value chains and the types of governance of these value chains will not be elaborated on here, thereby allowing readers to focus on the particularities of the labor process. Using the data collected from the research fieldwork, three main aspects of the content moderation labor process are highlighted: the recruitment process, the organization of work, and the conditions of work. These and other aspects of the moderation labor process have been expanded in further detail by Ahmad and Krzywdzinski (2022).

In terms of recruitment, the suppliers undertake most of the processes according to the SLAs, which specify the project details, including the number of workers to be hired by the supplier company. Depending on the agreements established between the two parties, some social media companies could directly participate in the recruitment process. The skills required for this work are mostly generic and allow applications from a diverse range of backgrounds, such as engineering and technology, media, and communications, management studies, and others. Opacity around content moderation production already starts from the recruitment process, where the moderators are required to sign non-disclosure agreements, thereby disallowing them from disclosing any details about the client and work process to a third party. Many of those who are selected and have agreed to exchange their labor for low wages and few benefits are "freshers," whose first job is content moderation. Overlooking the lack of work information provided to them, the moderators noted that they were

---

8   The analysis on the outsourcing relationships presented here and the resulting power asymmetries, is informed by an extensive literature on global value chains, most notably by Gereffi et al. (2005) and Ponte and Sturgeon (2014).

attracted to the possibility of working for *global brands* (popular social media companies) and saw it as their entry job into the IT sector.

The aspect of work organization can be explained according to the different types of content moderation. *Proactive moderation* before the content is published on the platform and *reactive moderation* after the content is published on the platforms are the two categories provided by Grimmelmann (2015) to explain the segmentation of the global content moderation market. A crucial point to note across both of these moderation types is the deployment of technical resources by social media companies. Extreme content, such as child sexual abuse and non-consensual porn, which impedes the public image of the company above the broad threshold, requires automatic detection before or very quickly after it is published on the social media platform. Interviews with both the content moderators and the representatives from the supplier companies revealed that such content does not enter the manual queues. Through the last years, many big social media companies have invested in or acquired the use of automated technologies to proactively moderate content. However, noting the large scale of content generated by users on their platforms, proactive moderation can be difficult. Reactive moderation depends on the users or third parties flagging or reporting content on the platform, and the content is sent to both automated and manual moderation processes.

Depending on the requirements of the social media companies, suppliers invest in basic filters or advanced technology, which constitutes the first part of the moderation process. Thereafter, content that has not been moderated by automated technology enters the queues of the moderators. These content queues can be identified as hate speech, spam, and others that are assigned to the moderators on mostly an arbitrary basis, following a mandatory training period. Depending on the terms of the SLAs, moderators review the user-generated content and make prescribed decisions according to the policies of the respective social media platforms.

The decision-making capabilities of the moderators vary from one moderation value chain to the other, where, on the one hand, the moderators are allowed to delete the content and even ban the user, and on the other, the moderators are allowed to simply tag the flagged content with the respective policies. Again, depending on the arrangement made between the social media company and the supplier, there exist other teams of quality analysts and team leaders that constitute fewer members and are higher up in the process hierarchy. Their work

comprises controlling the performance of moderators and may even include the task of making final decisions on the already tagged content by the lower-level of moderators. The work of content moderators is organized through moderation software and assistive technologies, which are either developed in-house by the social media company or have to adhere to stringent standards.

The organization of work has a multidimensional impact on the working conditions of the moderators. First, the content moderation work process is highly controlled, with specific targets assigned to each moderator every month, depending on their content queues, content format (videos, text, images, etc.), and team size. If they are unable to complete their targets on time, the management at the supplier company penalizes the moderators using gradually-depraving disciplinary measures. In the beginning, they are issued statutory warnings, following which they are shifted to elementary levels of content moderation work, or a simpler project. Granting all these steps, if the moderators are still unable to improve their performance, they are eventually expelled from the supplier company and are required to serve their notice period.[9] This creates a lot of psychological stress for the moderators and intersects with other reasons for resentment against the management, including low wages, long working hours, work-shifts[10], and lack of skill development.

The second trying element of this work is the distinct characteristic of the user content on which moderators have to review, tag, and or make decisions on. Content involving violence, assault, animal abuse, and other distressing material is visible to the content moderators, although the frequency of its visibility depends on their content queues. This means that queues with content on hate speech, violence, and nudity, etc., have a higher prevalence of distressing content than other queues, especially in the electronic commerce (e-commerce) section (such as Facebook Marketplace etc.). Regardless of the rate of occurrence

---

9   The notice period usually spans between one to three months and allows the moderator to apply for another project in the respective supplier firm. While their employment contract is still valid during the notice period, they are not paid their usual wages. Depending on the policies of the supplier company, the management might only support the health insurance of the moderator and even their families, which amounts to a small sum.

10   Content moderation service constitutes a 24-hour work cycle with three to four shifts running throughout day and night.

of distressing content, conversations with moderators during this research revealed that watching harmful content can have a lasting impact on the mental health of the respective moderators.

Considering the deplorable working conditions presented here, the reader might expect the emergence of collective resistance by the content moderators, especially since the Indian IT sector provides us with increasing examples of unionizing activities.[11] Instead of engaging in explicit forms of resistance, many of those who participated in this study exercised resilience and were of the view that they had to adapt to watching distressing content if they wanted to continue working in the content moderation process. Further, some echoed the opinion (by the management at the supplier companies) that their work was necessary to "guard the world against harmful content on social media." In terms of negotiations for wages and skill development, moderators approached the management individually, hoping to succeed on the basis of their personal relationships. However, the supplier management was often dismissive of these demands. Correspondingly, the social media companies played no role in managing conflicts between the moderator and the supplier company. Lacking possibilities for better career opportunities at the supplier company and the non-likelihood of employment at the respective social media companies (which they had initially aspired for), moderators design their own "career staircases" (James & Vira, 2012, p. 3; Ahmad & Krzywdzinski, 2022, p. 90) across the expanding labor market for content moderation in India.[12]

---

11   Indian IT trade unions, such as Union for IT-enabled Services (UNITES) professionals and Forum for IT Employees (FITE) have been formed in the last few years, mostly as a response to rising layoffs in the sector. Further, the IT and IT-enabled services sector is increasingly becoming a focus of interest for many central trade unions in the country.

12   There is a growing content moderation market in India with domestic and regional social media companies, including Chinese companies, sourcing content moderation services from the Indian suppliers.

5   Essential to social media but invisible to the world: Turning the spotlight on content moderation labor

Against the background of increasing public pressure to regulate social media platforms, this chapter presses for additional attention to the production processes of content moderation. This includes identifying the outsourcing practices that social media companies design to obtain content moderation services for their platforms. To this end, this chapter focuses on the labor process and the resulting working conditions of the moderators. The focus of existing scholarship on the Global North is expanded here to include India, where a growing number of content moderators are located. Much of this narrow focus can be attributed to the hidden outsourcing practices of content moderation, which veil the high power asymmetries between social media companies based in the Global North and content moderation supplier companies located in India.

The relationship between the two stakeholders has important consequences for the content moderation labor process. Social media companies outsource content moderation work to suppliers in India on a project basis and set standards for moderation policies, technology, and other product-related tasks. The companies in India are mostly tasked with employing the content moderators, controlling their performance and managing their wages, skill development, and other human-related aspects. The resulting labor processes have been described in this chapter under three main parameters: the recruitment process, organization of work, and working conditions. Lack of explicit forms of collective resistance by the content moderators is accompanied by their resilience and individual strategies for change.

This essay does not seek to provide an overarching picture of the outsourced content moderation practice to India. For starters, there is no single practice of content moderation outsourcing that can be delineated here. Instead, there are different content moderation value chains taking shape through agreements formed between social media companies and supplier companies in India. The governance of these value chains differs, leading to different levels of coordination mechanisms and power asymmetries. We can however note that most of the standards of content moderation are set by social media companies, thereby leaving the suppliers at less powerful positions and the moderators with even lower control over their labor. Further research is required to determine if there are more stakeholders involved in these chains, the mobility of workers across these

chains, and potential new forms of resistance. Additionally, domestic and international public policies must be aimed at improving the working conditions of moderators who supply commonly used social media platforms with essential labor.

Sana Ahmad is a doctoral candidate at the Freie Universität Berlin and an associate researcher at the Weizenbaum Institute for the Networked Society in Berlin, Germany.

References

Agence France-Presse (2017). Social media sites face heavy hate speech fines under German proposal. *The Guardian.* https://theguardian.com/media/2017/mar/14/social-media-hate-speech-fines-germany-heiko-maas-facebook

Ahmad, S., & Krzywdzinski, M. (2022). Moderating in obscurity: How Indian content moderators work in global content moderation value chains. In M. Graham & F. Ferrari (Eds.), *Digital work in the planetary market* (pp. 77–95). International Development Research Centre and MIT Press.

Batt, R., Doellgast, V., Kwon, H., Nopany, M., Nopany, P., & da Costa, A. (2005). The Indian call centre industry: National benchmarking report strategy, HR practices, & performance. *CAHRS Working Paper Series.* https://hdl.handle.net/1813/77401

boyd, d. m., & Ellison, N. B. (2007). Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication*, *13*(1), 210–230. https://doi.org/10.1111/j.1083-6101.2007.00393.x

Buni, C., & Chemaly, S. (2016). The secret rules of the internet. *The Verge.* https://theverge.com/2016/4/13/11387934/internet-moderatorhistory-youtube-facebook-reddit-censorship-free-speech

Burgess, J., Marwick, A., & Poell, T. (Eds.) (2018). *The Sage handbook of social media.* SAGE. https://doi.org/10.4135/9781473984066

Chen, A. (2012). Inside Facebook's outsourced anti-porn and Gore brigade, where 'camel toes' are more offensive than 'crushed heads. *Gawker.* https://gawker.com/5885714/inside-facebooks-outsourced-anti-porn-and-gore-brigade-where-camel-toes-are-more-offensive-than-crushed-heads

Chaudhuri, P., Chatterjee, A., & Verma, V. (2014). Guardians of the internet. *The Telegraph India.* https://www.telegraphindia.com/7-days/guardians-of-the-internet/cid/1669422

Dyer-Witheford, N. (2010). Digital labor, species-becoming and the global worker. *Ephemera: Theory and Politics in Organization*, *10*(3), 484–503.

Fuchs, C. (2010). Labor in informational capitalism and on the internet. *The Information Society*, *26*(3), 179–196. https://doi.org/10.1080/01972241003712215

Fuchs, C. (2014). *Digital Labour and Karl Marx*. Routledge.

Gandini, A. (2020). Digital labor: An empty signifier? *Media, Culture & Society, 43*(2), 369–380. https://doi.org/10.1177/0163443720948018

Gereffi, G., Humphrey, J., & Sturgeon. T. (2005). The governance of global value chains. *Review of International Political Economy*, *12*(1), 78–104. https://doi.org/10.1080/09692290500049805

Gillespie, T. (2018a). *Custodians of the Internet: Platforms, content moderation, and the hidden decisions that shape social media.* Yale University Press.

Gillespie, T. (2018b). Regulation of and by platforms. In J. Burgess, A. Marwick, & T. Poell, (Eds.), *The SAGE handbook of social media* (pp. 254–278). SAGE.

Gottlieb, B., & Karatzogianni, A. (2018). *Digital materialism: Origins, philosophies, prospects.* Emerald Group Publishing.

Grimmelmann, J. (2015). The virtues of moderation. *Yale Journal of Law and Technology*, *17*(1), 42–109.

Helmond, A. (2015). The platformization of the web: Making web data platform ready. *Social Media + Society*, *1*(2). https://doi.org/10.1177/2056305115603080

Herman, A. (2014). Production, consumption, and labour in the social media mode of communication and production. In J. Hunsinger & T. Senft (Eds.), *The social media handbook* (pp. 30–44). Routledge.

James, A., & Vira, B. (2012). Labour geographies of India's new service economy. *Journal of Economic Geography*, *12*(4), 841–875. https://doi.org/10.1093/jeg/lbs008

Medeiros, B. (2017). Platform (non-)intervention and the "marketplace" paradigm for speech regulation. *Social Media + Society*, *3*(1), 1–10. https://doi.org/10.1177/2056305117691997

Ponte, S., & Sturgeon. T. (2014). Explaining governance in global value chains: A modular theory-building effort. *Review of International Political Economy*, *21*(1), 195–223. https://doi.org/10.1080/09692290.2013.809596

Roberts, S. T. (2016). Digital refuse: Canadian garbage, commercial content moderation and the global circulation of social media's waste. *Wi: Journal of Mobile Media*, *10*(1), 1–18.

Roberts, S. T. (2017). *Content moderation.* In L. A. Schintler & C. L. McNeely, (Eds.) Encyclopedia of big data (C: 1–4). Springer.

Roberts, S. T. (2019). *Behind the screen: Content moderation in the shadows of social media.* Yale University Press.

Stone, B. (2010). Policing the web's lurid precincts. *The New York Times.* https://www.nytimes.com/2010/07/19/technology/19screen.html

Wong, J. C. (2021). Banning Trump won't fix social media: 10 ideas to rebuild our broken internet – by experts. *The Guardian.* https://theguardian.com/media/2021/jan/16/how-to-fix-social-media-trump-ban-free-speech