IGNORED AND DELETED: UNDERSTANDING CONTENT MODERATORS AS RACIALIZED MEDIA OF SOCIAL NETWORK SERVICES

SEUNG WOO BAEK

A THESIS SUBMITTED TO THE FACULTY OF GRADUATE STUDIES IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF ARTS

GRADUATE PROGRAM IN COMMUNICATION AND CULTURE

YORK UNIVERSITY

TORONTO, ONTARIO

AUGUST 2022

© SEUNG WOO BAEK, 2022

ABSTRACT

This thesis investigates how Facebook moderates its social media platform and mediates content flow by employing subcontracted Filipino workers as a form of racialized media filter. Scholarships on social media networks have often focused on the contents that flow through its networks, rather than the material and historical make-up of the infrastructure that enables said circulation. The thesis seeks to highlight the colonial and racist logic that undergird commercialized content moderation and its practice of global labour outsourcing that seeks to meet the Western social media and tech companies' demand for cheap, fast, and available labour. The research looks to the history of transcontinental railway and its usage of Chinese migrant labour as a parallel media history and to Armond R. Towns' "Black mediality" as a conceptual framework that helps illustrate the colonial mode of racialization inherent in contemporary network of social media.

DEDICATION

To those who stand in the middle.

ACKNOWLEDGEMENT

I am thankful to have the time and space here to acknowledge that this thesis is a product of, and my personal response to, the connections that I have been able to forge and nurture in the past years. A significant part of this research has come together during a harrowing time of distancing and delayed connections; recognizing that, I am all the more thankful to those who supported and encouraged me to dedicate my attention to acts of learning and exploring and thinking.

I would like to thank Professor Greg Elmer for his supervision and guidance in helping me shape my interest into a form of knowledge that I am proud to share with the wider community of the learning and research. When I started this study, I only had a handful of questions, but now, after two years, I have more questions than ever before to accompany this thesis. This seems a great progress, and I could not have done it without his support and critique.

I would also like to express my gratitude to Professors John Shiga, Rianka Singh, and Susan Ingram for sharing their expertise, care, and time, as members of my thesis committee. What I have to offer in this thesis would not have been possible without your generous inputs.

To my family and friends, thank you for your love and patience, care and interest, concern and celebration. Every single page of this thesis is fuller for your presence in my life. Thank you, especially, for asking all along how my thesis was going. I am glad to say it is done.

I would lastly like to acknowledge the support of the Social Sciences and Humanities Research Council (SSHRC), without whose financial support I would not have been able to complete this thesis.

TABLE OF CONTENTS

ABSTRACT II

DEDICATION III

ACKNOWLEDGEMENT IV

TABLE OF CONTENTS V

CHAPTER 1. RACIALIZED COMMERCIAL CONTENT MODERATION 1

Introduction: The stories of The Cleaners (2018) 1
Definition: Content moderation as CCM 8
Extended definition: Amalgam that produces knowledge 12
Object and Method: Focalizing through the site of Facebook 15
Gap: Mediatization of human bodies for trauma-sensitivity 18
Select research questions 21
Lacks and limits 22
Roadmap 25

CHAPTER 2. NETWORK ECONOMY, OF HISTORY AND PRESENT 28

Missing bodies in our discussion of "networked and digital economy" 28
"Black mediality" in our networks 32
Understanding Colonial mediality 37
Reading the history of transcontinental railway through the framework of Colonial mediality 39
Reading commercialized content moderation through the framework of Colonial mediality 50
Conclusion 66

CHAPTER 3. AI-SOLUTIONISM AND RACIALIZED HUMAN-MACHINE PARTS 69

New bodies in our discussion of "networked and digital economy" 69
The growing problem called "social media" 71
The solution of more: bodies and machines 72
The Real AIs in the solution 78
Conclusion 88

CHAPTER 4. MEDIA OF THE FUTURE 92

The next thing 92
Future looking 93
Race as technology 96
The resistance 98
What next? 103

REFERENCES 106

CHAPTER 1. RACIALIZED COMMERCIAL CONTENT MODERATION

> Is not the landscape, every glimpse of which hath a grandeur, a face of him? Yet this may show us what discord is between man and nature, for you cannot freely admire a noble landscape, if laborers are digging in the field hard by.
>
> -- Ralph Waldo Emerson, *Nature*

> The 'life' of the message purchased with the death of the messenger; the messenger sacrificed through the delivery of the message; Is there a connection between being a messenger and being a sacrifice? (Krämer, 2015, p. 37)

Introduction: The stories of The Cleaners (2018)

The Cleaners (Block & Riesewieck, 2018) opens with a bustling scene of night crowds as they navigate the booming streets of Manila, Philippines. The brightly lit people weave in between all-glass high-rises whose many office floors are dark yet remain abuzz with work into the early hours. The documentary's main stories unfold in these dark rooms, whose rows of computer screens light up parallel rows of faces staring at them. Screens flicker through an endless stream of images and videos that carry scenes of violence or pornographic details, and they are accompanied by muffled sounds of crying and screaming. Seated in front of the computers, the content moderators, the documentary's main subjects, review online content. They must, in a matter of seconds, determine whether the content should be deleted from the platforms they moderate, or permitted to stay. The moderators speak their decision out loud – "ignore" or "delete" – and then click on to finalize their decision.

Punctuated by excerpts of email communications the filmmakers had with content moderators, the documentary follows and hears from, among others, a select number of former and current content moderators working in Manila, Philippines. Some show their faces, but many more do not, and we do not learn their names. But we hear of their experiences as employees of subcontracting firms working at the behest of Western tech companies and social media giants

like Google and Facebook. These are the moderators that undergird the social media platforms that billions use around the world daily.

In the documentary, one moderator talks of witnessing child pornography that were uploaded onto the platform, and of how when they reported to their team leader that they simply could not continue on with their work because it was too much, they were met with a cold shoulder: "It's your job," they were told. "You signed a contract" (Block & Riesewieck, 2018). Another talks of watching "hundreds of beheadings" over the course of their career as a content moderator, so many that at one point they could identify what weapon was used to commit the act by just looking at a still photo of a mutilated body (Block & Riesewieck, 2018).

The work they do is brutal and ceaseless, and it is often done from afar. One content moderator's work (the contents that they had to screen) at the time of filming was coming mostly from the United States, around the "time of the winning of Donald Trump"; the moderator talks about reviewing posted "slurs against immigrants, against different races, like Mexicans and Blacks" (Block & Riesewieck, 2018). One moderator recalls a time when he had to watch a livestream of someone attempting self-harm, and how he was unable to shut down the stream as long as the person had not "actually committed suicide" (Block & Riesewieck, 2018). So, the moderator sat at his station and watched, waiting for the person to take the last drastic action, one that would trigger the final link in the chain of moderation decisions. Another, through an email correspondence, tells of a moderator who "specialized on [sic] self-harm live videos" and how, due to the incredible stress caused by this work, they ended up committing suicide (Block & Riesewieck, 2018). The constant stress and exposure to vicarious trauma, the documentary tells, have led many former content moderators to quit their jobs. Many entered these jobs considering them to be high-paying tech jobs, a form of "white-collar" position, a way out of poverty or

simply a desirable way to make a living. These jobs had instead turned into a lasting source of mental and physical health burdens.

Such devastating testimonials are unfortunately not rare in the industry of content moderation and, by extension, in the business practices of major social media companies, including the GAFA companies.[1] They collectively provide an revealing narrative of what is at the foundation of social media platforms: someone somewhere is doing the necessary work of sanitizing our social media platforms so that we can “safely” share content and communicate with each other. When we seek to create forums of discussion and build communities, in the middle of those social connections, there is a person doing the damaging work of moderation, often in secrecy, away from view, and at an incredible personal cost.

My research into the industry practice of content moderation began with *The Cleaner*. However, since then, I have come across many other stories, collected by journalists and scholars, that corroborate one another’s work and cumulatively painted a picture of an industry that outsourced a gruesome form of labour to the “global south.” Through their outsourcing tactics, the predominantly Western tech companies could operate their social media platforms by acquiring cheap labour, at-scale, and by keeping their operations at a legal arms-length, they could discard the individuals that they harmed in this process.

It is important to note that my usage of the term “global south,” in this sense, does not merely focus on geographical locations that have been impacted heavily by the colonial violence of the Western imperial expansion (like that of the Philippines and its history as an American

[1] Borrowing from Alan B. Albarran’s definition of the social media industry, I am looking to Facebook as a member of a group of American Big Tech companies that offer products and services to buyers in various markets. The seller group may involve other technology service companies that comprise GAFA, referring to Google (now Alphabet), Amazon, Apple, as well as Facebook (now Meta), while the buyers are “a combination of consumers who create their sites with social media technologies and advertisers who purchase access to these audiences with the hopes of selling their own products and services” (Albarran, 2013, p. 2). There is a notable relationship between social media services/technology firms and its user-base as consumers, even if such services (and platform access) are being proffered for “free-of-charge.”

colony). For me, “global south” speaks not only to the power differential that such histories have created, but also to the strategy of power that discursively and practically imagines the global south and its peoples as raw material and a ready-to-use labour pool, as well as a destination of digital wastes and harmful externalities of imperialist capitalism.

While “global south” is a term enmeshed in ongoing debates about its intellectual lineage and efficacy in describing today’s world, I am in part looking to invoke Vijay Prashad’s articulation of “a new geography of production” that sees “relocation of production processes from the First to the Third World, thus allowing corporations to take advantage of wage differentials and of significantly fewer restrictions on environmental and labour standards” (Prashad, 2012; Sajed, 2020). I thus engage with the question of how commercialized content moderation, and its practice of global outsourcing, continue to exercise colonial and racist logic to meet Western social media and tech companies’ demands for cheap, fast, and available labour, just as it has done in the past with other forms of information networks, such as railways. I will return to this specific connection in the second chapter of the thesis.

In recent years, content moderation has become a more prominent topic in discussions about digital spaces and social media platforms, in part thanks to the 2016 US election scandal surrounding the allegations of Russian interference via platforms like Facebook; the scandal raised concerns about social media platforms’ potential for misuse by bad-faith actors who could spread misinformation and initiate targeted propaganda campaigns throughout the broader public. The concerns largely coalesced into questions about how social media platforms and Big Tech companies are (or otherwise not) monitoring such campaigns and moderating dangerous contents off their platforms. However, much of the attention, when available, has focused on content moderators as an unfortunate but consequential, and even necessary, presence in the

operation of social media platforms and the internet economy[2] (Gillespie, 2020; Steiger et al., 2021). However, rather than seeing content moderators as playing a secondary, almost an incidental, role of filling in the gap of an otherwise well-functioning social media platform technology, this research looks at how the role of content moderator is actually made foundational to the mediation of social media platforms and, by extension, a significant portion of our contemporary infrastructure of digital communication and economic exchange through the violent racialization of the bodies of those that take on this role and do this work. Part of this emphasis also includes understanding that the flow of online digital content is not such a neutral activity, and that our understanding of social media networks must examine their racial and technological make-up and how such materiality of the infrastructure impacts the mediation of said content (whether harmful or not), especially given the violence against a racialized labour force that these networks require for their routine functioning.

This introductory chapter opens the thesis by outlining and contextualizing the stories of content moderators and the work of moderating social media platforms visited by billions of people daily. It outlines and defines the practice of commercialized content moderation and explores related concepts of networked and digital economies. In doing so, this chapter begins the illustration of how the global practice of outsourcing is an embedded part of the knowledge economy enabled by social media and its content moderation practices. The chapter provides a literature review of the scholarly and journalistic coverage of the social media industry's exploitative practices relating to content moderation. It closes by outlining certain limitations of

[2] The term "internet economy" is taken from Steiger et al. (2021) and describes a new form and domain of economic exchange and market activity made possible through the emergence of internet and internet-enabled connections. In this sense, the term points to content moderation as a novel function within the ecosystem of social media networks and as a unique labour need that came about due to the advent of internet.

the research, and by presenting the set of research questions and methodology that shape the overall research.

Part of my contextualization of this contemporary phenomenon includes looking to the history of the infrastructure of communication and media as paralleling the stories of the content moderators and their advocates. In specific, I look to the case study of the construction of the transcontinental railway in the late 1800s, which made significant use of migrant Chinese labourers to construct much of its nation-spanning transportation network. The Chinese labourers were often given the most dangerous tasks of the railway building, tasks with a high death toll. This violent history, I argue, continues today in racializing certain bodies in a process that discounts their labour and devalues their humanity, all in order to materialize communication networks and infrastructures. In the context of content moderation practices today, this violent racialization has manifested in psychologically damaging ways, in the form of monotonously demanding labour which has been outsourced to global south – namely to Philippines. Both infrastructures of connectivity (transportation and communication) materially and discursively constructed and networked America's imperial economy. In comparing these two histories, I hope to help illuminate how the necessity of content moderation and its outsourcing to places like Manila are not a natural or rational outcome of operating an internet-enabled social media network. Rather, I argue that the modern practice of content moderation necessitates racializing a class of humans so that they are "opened up" to traumas (i.e., enacting violence against and through) with the ultimate goal that racialized bodies can be called to serve as a form of media, one that can filter out harmful contents from the platforms that they labour for, so that its users can digitally mediate and communicate across great distances and through betwixt temporality in the mode that we know as social media.

Of course, the specificity of racialization that affected the Chinese migrant labourers of 1800s and the content moderators of Manila, or elsewhere in the Philippines, today is different. In my comparative reading of the histories of communication infrastructures, I endeavour to indicate and be cognizant of how their dispossession of bodies, which took place in service of the materialization of the infrastructure, occurred under varying modes of control. However, even though the type and texture of labour that the Chinese migrant labourers of 1800s and the contemporary content moderators in the Philippines differ, I point out that the ways in which their bodies are mediatized, utilized, and then discarded are core aspects of how they formulate the infrastructure of communication then and now. Sitting as the extension of the colonial history of American wealth and development, the violent racialization of the content moderation work and its globalized labour force has become a part of the very infrastructure of modern communication, the digital economy, and social media platforms.

Lastly, I seek to theoretically situate this reading of communication infrastructures by looking to the works of media philosophy and critical race scholars like Armond R. Towns and Beth Coleman, who have given us tools and frameworks of thought like “Black mediality” and “race as technology.” Towns’s work on “Black mediality” grounds my approach to understanding how racialization of bodies have been foundational to the history, philosophy, and scholarship surrounding media and technology. I look to apply Towns’s framework to my historical readings and to my examination of the content moderation industry, thus extending the framework of “Black mediality” into one of Colonial mediality responding to the colonial and imperial history of American network infrastructures.

Definition: Content moderation as CCM

It is important to understand that stories from *The Cleaners* (Block & Riesewieck, 2018) speak of a very specific configuration of the labour, a slice that is part of a bigger digital mediation process (i.e., of sending, screening, and receiving digital communication) labelled "content moderation." To better understand this slice, it would benefit us to understand how the overall scheme of content moderation has been discussed so far.

Sarah T. Roberts is one of the more prominent scholarly voices that discuss the matter of content moderation today, and is also the person who coined the term "Commercialized Content Moderation" (CCM) to describe the "organized practice of screening user-generated content (UGC) posted to Internet sites, social media and other online outlets, in order to determine the appropriateness of the content for a given site, locality, or jurisdiction" (Roberts, 2017). Apropos of social media platforms in particular, Roberts points out that the content moderators are said to be employed to maintain spaces that are "more palatable, accessible, and inviting" and that their labour elicits "more user participation" as a result (Roberts, 2019, p. 39). While content moderation can be undertaken by volunteers or users themselves, CCM points to the commercial context in which individuals or firms provide content moderation services and receive renumeration (Roberts, 2017).

Indeed, while content moderation as a practice has been a part of internet-enabled connection since the days of bulletin board system (BBS), often exercised as a volunteer role, the shift to the commercial formulation of the work marks a significant turning point in the broader context of digitally-enabled communication (Ackermann, 2020; Roberts, 2017). Roberts notes that as internet adoption gained strong traction, mainstream platforms' desire to "control the UGC that they host and disseminate" grew as well (Roberts, 2017, p. 2). In the midst of this

growth, “newspapers and other news media outlets” became the early adopters of CCM, which they enlisted in an effort to “combat what they viewed as the misappropriation of the comments spaces,” which they noticed “often devolved into unreadable spaces filled with invective, racist and sexist diatribes, name-calling and irrelevant posting” (Roberts, 2017, p. 2). This need for control over UGC and the web environment of platforms continued alongside the rise of social media networks like Friendster and MySpace.

The individuals so employed in the commercial context are often “human actors who rely upon their own linguistic and cultural knowledge and competencies to make decisions about UGC’s appropriateness for a given site or platforms,” and today they are overwhelmingly represented (or employed) by Business Process Outsourcing (BPO) companies who serve as a middle-man entities between the moderators and the tech platforms, with some being employed through micro-labour sites like the Amazon Mechanical Turk (MTurk) as contractual employees.

While it is important to acknowledge that content and community moderation as a practice has existed (and indeed continues to exist) in various forms since the early days of Internet (Ackermann, 2020; Carlson & Cousineau, 2020; Jhaver et al., 2019), often involving “voluntary and self-organized capacity” (Roberts, 2014, p. 68), Roberts’s definition is notable for the distinction that it makes between a voluntary digitally-mediated sociality and a commercialized one. This distinction provides a rich starting point for our discussion of content moderation as a globally industrialized practice that binds its moderators in a contractual relationship to their distant employers and the platforms that they moderate. As well as mediating platform sociality, such as what is permissible in digital spaces, commercialized content moderators become an important part of a production cycle of analytical work that creates “immaterial products, such as knowledge, information, communication, a relationship, or

an emotional response" that circulate through a given platform (Roberts, 2019, p. 39). This very circulation undergirds the incredible valuation that many social media and tech companies enjoy, both in their actual capacity to mediate content and collect sellable user data, and in their more speculative valuation and capital-seeking behaviours (Elmer, 2019). To either of these ends, content moderators are indeed considered a crucial aspect that enables the smooth operation of media platforms and sites that seek to attract and maintain high volume and intensity (e.g., frequency) of user engagement. In this understanding, the content moderators can be said to be employed by businesses and tech platforms not only in their function as screener of contents that "curate" the "character" of the platform or the digital spaces, but also as an extended and consequential means of managing user experience, generating capital value, and as a "means to protect the firm from liability, negative publicity" (Roberts, 2017). Their presence further implicates them in global companies' legal strategies to ensure social media platforms' ongoing operation in adherence to various local, national, and supranational jurisdictional norms and regulations, as these companies seek to make their platforms accessible to an international userbase, in part by connecting diverse networks and functions.

Employment of content moderators at a commercial scale to safeguard corporate interests became all the more apparent in the aftermath of the 2016 US election interference allegations, which gained significant political currency and, coupled with the backdrop of the Cambridge Analytica scandal of 2018, became one of the precipitating circumstances that drew significant media attention to social media's role – especially that of Facebook – in allegedly enabling and permitting said interference (Roberts, 2019, p. 24). When pressed by congressional to respond to the allegations, Facebook responded by increasing the scale of content moderation, much of

which was outsourced to places like Philippines and India (Arsht & Etcovitch, 2018; Dwoskin, 2019; Jardine, 2019).

Much of the scholarly attention that has touched on content moderation has debated issues of freedom of speech (Ganesh & Bright, 2020; Kaye, 2018), harmful speech moderation (Einwiller & Kim, 2020; Siapera & Viejo-Otero, 2021), and related issues of platform governance, often further zeroing in on the platforms' inability to adequately moderate and, thus, govern its platforms and user base (Fagan, 2020). The contemporary struggle against the global pandemic added another layer of context that contributed to a rise of misinformation threatened public health, incited racially motivated violence, and overall exacerbated many extant societal tensions negatively affecting the general sense of public safety and private well-being. Social media platforms were often at the centre of such discussions, noted for their inability to, on the hand, limit the rise of misinformation, and on the other, for their perceived over-reach limiting free speech in an attempt to curtail the spread of harmful content.

In this regard, scholars like Sarah T. Roberts and Tarleton Gillespie introduce a meaningful thesis that content moderation (and thus discussion of content moderation) at scale requires "immense human resources" to the contemporary consideration of the role of social media (2018, p. 115). The way they have endeavoured to typify and categorize a "diagrammatic" (Elmer, 2004, pp. 43–44) approach to understanding content moderation as being performed on a social, material, and technological level have added to the growing understanding of how the CM labour is distributed unevenly and as an embedded process of modern communication. This growing understanding of content moderation casts significant doubt on tech companies' repeated promise to reduce human harm and cost by means of scaling their operation or by future automation (Gillespie, 2020). Other scholarly works on content moderation's impact on

information relevance and quality (Link et al., 2016; Mulla & Palave, 2016; Veglis, 2014), user agency that circumvents certain platform affordances or generalized web access as they relate to content moderation (Chancellor et al., 2016; T. M. Chen & Wang, 2010), or other user-centric research on user-generated contents and user expectations of moderated platforms (Greis et al., 2014; Myers West, 2018), including user-directed content moderation (Ackermann, 2020; Lampe & Resnick, 2004; Momeni, 2012), have further expanded the on-going discussion on how content moderation is done and towards what end.

Extended definition: Amalgam that produces knowledge

Emerging out of this body of research is the understanding and acknowledgement that CCM is an incredibly layered and interconnected activity that requires an amalgam of processes, knowledges, inputs, and relations, including user engagement and automation apparatuses. In this amalgam, the contractually (i.e., commercially) employed moderators occupy a specifically classed position within the infrastructure of social media and networked communication. While conventional wisdom focuses on the specific moment or the mode of moderation (i.e., making the decision to "delete" or "ignore" the content) as the emblematic image of content moderation, the overall schema of content moderation is far more embedded and expansive, reaching into the broader infrastructure of global telecommunication and labour outsourcing (Barrett, 2020).

To better understand this layered complexity and diversity, let us consider for a moment how the process of moderation begins. A content can enter the moderation pipeline, for example, by users using the "flagging" functionality to object to certain contents that they are seeing on the platform. Contents can also be flagged by external parties or legal representatives who make the request for deletion directly to the platforms' business entities based on allegations of copyright infringement or intellectual property theft. At times, law enforcement and other

governmental actors may approach the platforms and its business entities with requests for moderation, which may take the form of pre-screened lists that the platforms are asked to apply singularly or to integrate into their daily operations (Biddle, 2021). Automation and automated systems also participate in this process, algorithmically sorting contents in conjunction with or even in absence of human participation, albeit to varying degrees of success (Scott & Kayali, 2020). Moderation, beyond being a straightforward process of screening and deleting contents, can also involve processes of annotation and contextualization, which seeks to provide additional or clarifying contexts to contents that do not meet the criteria for deletion but are nonetheless deemed problematic; various COVID-19-related contents, including sales of masks and policy advertisements, were restricted in their distribution and were required to include specific disclaimers indicating who sponsored the contents on Facebook (Facebook, 2021).

However, the responsibility of determining what contents are moderated and how, as well as how such contents should be displayed and contextualized does not solely rest with moderators. Programmers, engineers, and policy specialists are often employed directly by the platforms themselves, at a much higher rate of pay, to be in charge of constructing these normative codes and documents that direct, in part, the content moderation process: "Perception of work importance is often focused on required skills rather than the necessity of the work itself. While skilled and creative IT work is often respected and well-compensated, relatively unskilled and rote data processing is typically afforded less stature and reward" (Steiger et al., 2021). These engineers and policy experts, who also enjoy more robust labour protection and health benefits, are tasked with drafting the guidelines and operational manuals that are then given to content moderators (often via contracting BPOs) who are tasked with interpreting and applying the manuals to the best of their abilities. While the process of "reading" and "knowing" the

content (as being either permissible or not, at its base level) takes place with inputs from both classes of workers, their positionalities are drastically different and so is the way they are treated.

Adding to the weight of unfairness in this distinction is how nonlinear and hodgepodge the rules of content moderation can be:

> Graphic, violent images should be removed, with the exception of aborted fetuses. Hate speech is prohibited, but only for protected categories of people (specific races, religions, sexualities, nationalities), and "migrants" constitute only a quasi-protected category, so although dehumanizing statements about them should be removed, cursing at them, calling them thieves, and urging them to leave the country do not amount to hate speech" (Gillespie, 2018, p. 112).

This means that the work of reading, observing, interpreting, and making decisions on the contents' permissibility on a platform requires a breadth of knowledge that covers a range of domains, including legal, social, cultural, and linguistic, as well as the weight of personal judgement to apply them. Application of the guidelines can also never quite be a simple matter of applying what has been codified. These "codes" (i.e., guidelines or "community standards," in the case of Facebook), to make matters worse, often shift and change in real-time, reflecting the fluctuating legal boundaries and socio-cultural terrain that platforms operate within. So, the moderators must be able to stay current on these shifting codes and integrate them into their decision-making process as quickly as they are updated. Part of this knowledge production also includes an affective dimension of moderating content, especially as a confluence of what Rae Jereza refers to as "the banal and horrific together with the complex sensations they provoke" (2021, p. 930). Content moderators are called to utilize all of their bodily being to gauge a sense of cultural taste and appropriateness, as well as the emotional and moral reaction (i.e., sense of wrong) that such contents elicit. Jereza refers to this dimension of their labour as being a form of "care work that moderators practise on themselves and on each other," and the management of their affective response as a form of "affective labour" (Jereza, 2021, p. 930). All in all, this

complexity of content moderation labour is not borne out by its "relatively low-wage and low-status" compared to other jobs in tech (Roberts, 2017). Otherwise put, while both classes are meant to produce knowledge about the content and, by extension, about the platforms that they moderate, what they produce and know are not valued equally, nor are they done in the same way.

Object and Method: Focalizing through the site of Facebook

There are lots of social media platforms to choose from when discussing content moderation. However, Facebook stands out for its (1) scale, (2) primacy as one of the, if not *the*, social media platform of daily use today, and for the (3) abundance of extant research and media coverage. Especially, when it comes to the way in which commercialized content moderation is carried out by the platform, Facebook seems to enjoy the most available and explicit coverage. Given these considerations, I seek to conduct a close reading of the industrial process of content moderation from various vantage points, including the perspectives of moderators themselves, companies' business filings, and journalistic coverage. In part, this availability of narratives further allows me to specifically zero in on select BPO firms and their relationships with Facebook, helping to address one of the most difficult challenges of researching social media platforms: the opacity of their technology and their business operations. These relationships, I argue, benefit by mediatizing its employees, who are often found in parts of the world that bear the violent history of colonialism, which in turn interpellates them to participate in the grander globalized imperial economy of the information economy in 21st century.

In so carrying out the above noted close reading of social media platforms and their practices of commercialized content moderation, I have identified Facebook as one of the primary sites. There exists some throughline of content moderation as occupying the role of

mediating a sense of connected sociality or in providing certain centralization of specific set of mediating and censoring power (Ackermann, 2020; Roberts, 2019, p. 5), however the scale of content moderation required today far exceeds the limits of ad-hoc solutions that were and are available in community-based or user-volunteer reliant methods of moderation.

According to Statista, the total number of social media users around the globe is expected to reach over 4.4 billion individuals, representing approximately half the world's population, with East Asia and North America having the highest penetration rate at 71 and 69 percent, respectively (Statista, n.d.; Statista Research Department, 2021a). Servicing them are myriad media platforms that cover a wide range of functions and purposes, but social networking, which includes "blogs, forums, business networks, photo-sharing platforms, social gaming, microblogs, and chat apps," is chief among them (Statista, n.d.; Statista Research Department, 2021a). Even among the throng, Facebook emerges as the "biggest existing social network" with 2.8 billion monthly active users and a revenue of 86 billion US dollars, as of 2020. Facebook's supremacy, not only as the premier social media platform, but also as a corporate entity Meta,[3] is illustrated by the fact that the top five most popular social media networks worldwide (as of October 2021 and ranked by the number of active users) include three of Meta's platforms: Facebook, WhatsApp, and Instagram (Statista Research Department, 2021b; "The Facebook Company Is Now Meta," 2021). The sheer dominance of Facebook as the premier social media platform has also been verified by academic research (Alhabash & Ma, 2017). Alongside Google, Amazon, Apple, and Microsoft, Meta has become one of the "Big Five" American information technology companies that touch on a significant portion of humanity's daily digital activity.

[3] Previously known as Facebook, Inc., Meta Platforms, Inc. rebranded as of October 28, 2021.

Its grand scale of operation is also reflected in the way Facebook has responded to the rising calls for more stringent and effective moderation of its platforms. Facebook has reported that it has spent more than 5 billion dollars in "safety measures" in 2021, including the worldwide contracts that it holds for more than 15,000 content moderators (Perrigo, 2022a). Barrett notes that on Facebook alone, "more than three million items are reported on a daily basis by users and artificial intelligence screening systems as potentially warranting removal" (Barrett, 2020). And in the first quarter of 2020, for example, Facebook reportedly removed or covered (a form of moderation that involves providing a warning about the content without removing it directly) about 1.9 billion spam items, 1.7 billion fake accounts, and 107.5 million pieces of contents that included "adult nudity, violence, and dangerous organizations" (Barrett, 2020). Biggest slices of the "107.5 million pieces of content" pie included: "adult nudity and sexual activity" (39.5%), followed by "violent and graphic content" (25.5%), "dangerous organizations: terrorism and organized hate" (11%), "hate speech" (9.6%), "drugs and firearms" (9.3%), "child nudity and sexual exploitation of children" (8.6%), "bullying and harassment" (2.3%), and "suicide and self-injury" (1.7%) (Barrett, 2020). To moderate this torrent of contents, Meta has employed more than 15,000 workers, majority of whom are outsourced through third-party vendors and multinational, overseas BPO firms to moderate "Facebook's main platforms and its Instagram subsidiary" (Barrett, 2020). This process has, according to Facebook, allowed the platform to comprise a review team that is global and one that provides around-the-clock coverage through "20 sites around the world, where these teams can review content in over 50 languages" (Meta, 2021b).

Gap: Mediatization of human bodies for trauma-sensitivity

What is missing from this expansive discourse on content moderators and the moderation process is how trauma and racialization has become a necessary part of the process of knowledge production and its digital economy. Even those that acknowledge the human cost as an embedded part of content moderation conceptualize trauma as an incidental externality of its process and economy. In other words, even the critics of social media platforms often consider the harm experienced by its content moderators as an outcome of a process that inputs harmful content through its network. However, looking at content moderation from the vantage point of the history of networked economy allows us to understand how trauma does not simply happen to the moderators. Rather, we see that the industry and its history conceptualize certain classes of labouring bodies in an instrumental relation to the trauma and the traumatic content that they are made to sense, moderate, and censor. Other scholarship that focuses on the production cycle of digital infrastructure and its hardwares have observed how important it is to understand the global circulation of digital goods as being a racialized process that has its roots in colonial history (Oyedemi, 2019, 2020; Roberts, 2016).

Lisa Parks argues that media infrastructure should be considered in "an historical and intermediale sense" and that "media infrastructures demand a consideration of the ways that distribution processes have emerged, changed, and *been layered upon one another over time*" (italics mine) (Parks, 2015, p. 357). Nicole Starosielski, in her study of undersea cable networks, makes a similar observation of the "resolute materiality of network infrastructure and its entanglements with the turbulent histories of the Pacific" (Starosielski, 2015, p. xi). My reading of content moderation similarly seeks to highlight the historical overlay that goes into understanding how our digital infrastructure and circulation continues to allocate harmful

externalities on a colonial and racist map of the world, one that we inherited since the days of the transcontinental railway. By looking at these historical entanglements, my research focuses on the ways in which moderation of the globalized infrastructure of digital network (that mediates communication and produces economic value through knowledge labour) is constructed by the very act of traumatizing the racialized bodies. In this sense, the trauma begotten from their position within the globally networked social media platforms is not a consequence but is the very thing that makes their racialized labour valuable to the social media platforms, its parent companies, and its subcontracting business partners. Otherwise put, my research points to how racialization has meant that certain bodies are made available for labour precisely because they are imagined to be the right body to handle the violence and harm inherent to certain types of labour, not despite them. This kind of dehumanizing logic and discourse sits at the bedrock of modern infrastructures of communication. Racialized workers are not only working to construct media infrastructures, rather they are the mediating infrastructure (that mediates harm and violence) itself.

In dedicating a chapter of her book, "Commercial Content Moderation in Manila: The Dating App Datakeepers of Eastwod City," to the stories of content moderators in Manila, Philippines, Roberts (2019) accounts for the colonial past of America's imperial violence as priming the ground for modern-day digital infrastructure that has made outsourcing practices to global south locations the norm. However, solely reading the Philippines' "abundant, relatively inexpensive" workforce that is "already intimately familiar… with American norms, practices, and culture" as having been constructed through "more than a century of political, military, and cultural domination by the United States" and thus seeing the rise of BPOs as the latest trend to impact the industry of content moderation, which otherwise had its humble beginning in

voluntary platform mediation in-house, boutique solutions employed in the early days of social media firms, overlooks the more foundational and temporally complicated relationship that exists between the racialized bodies of the Philippines' content moderators, the networked infrastructure of communication and commerce, and the history of American imperialism that constructed its networked infrastructures as an infrastructure of imperial economy. While acknowledging the chronological development and evolution of commercialized content moderation as being a relatively modern industry, contextualizing its development in relation to other histories of telecommunication (and transportation) networks and infrastructure makes it possible to understand how the racialization of its labouring bodies, which often came from beyond American shores, was an integral part of its construction, maintenance, and success. In other words, it is important to observe that the internet economy of social media platforms was made possible precisely by the configuration and existence of racialized labourers that serve as its medium, and in particular as bodies that are able to sense *accurately, efficiently, and uniquely* the offending materials permeating its networks. The infrastructure of communication required human beings to decipher its noise from signal, and it racialized certain bodies (i.e., "opened up [the bodies and beings] for violence") to occupy that role (Towns, 2019).

Journalistic and documentarian sources have been quick to focus on the physical and psychological harm (i.e., vicarious trauma and Post-Traumatic Stress Disorder) caused by the industry. Investigative journalistic (A. Chen, 2014; Dwoskin, 2019; Gilbert, 2021; Newton, 2019, 2020) and documentarian (Block & Riesewieck, 2018; A. Chen & Cassidy, 2017; Rigaud, 2019) interest in the stories of content moderation, including testimonials from former and present moderators, have far exceeded scholarly interest in making explicit the brutal reality of CM work (Ahmad, 2019). These sources have helped to expand the discourse on content

moderation by including personal stories and profiles of the content moderators, most of whom are legally barred from speaking out themselves at the risk of social and personal harm, by non-disclosure agreements (NDA) that they are forced to sign as a part of their employment arrangement. This coverage has done the invaluable work of getting these stories out. However, even in these discussions, the focus at times can be on the gruesomeness of the details of the work and the sense of bizarreness of their labour, thus running the risk of exoticizing the problem and side-stepping the underlying logics that create the problem in the first place. Telling headlines include: “Trauma Floor” (Newton, 2019); “Bestiality, Stabbings, and Child Porn: Why Facebook Moderators Are Suing the Company for Trauma” (Gilbert, 2019); and “The Labourers Who Keep Dick Pics and Beheadings Out of Your Facebook Feed” (A. Chen, 2014). My research looks to these sources to help augment my reading and prioritize the lived experiences of the moderators, while staying centred on the relationship between “harm” and the racialized workforce’s role in mediating that harm, not only as a sensational consequence, but as the inherent and intended design of content moderator positions.

Select research questions

My research is guided by four major questions: (1) How does commercialized content moderation, and its practice of global outsourcing, continue to exercise a colonial and racist logic to meet Western social media and tech companies’ demand for cheap, fast, and available labour? (2) How can we gain a more nuanced understanding of this logic by looking to the history of the construction of transcontinental railway, another information network that made use of a racialized workforce (i.e. the migrant Chinese labourers) to construct its infrastructure? (3) How can we broaden our understanding of the materiality of digital communication infrastructures by looking to racialized bodies’ role in interfacing the amalgam of software, hardware, and

protocols of knowledge production? And lastly, (4) What possibilities for resistance are available to those interpellated into the industry of content moderation and made to serve as a form of media filter?

Lacks and limits

One of the concerns of this research is that it is difficult to describe a singular (or typical) working day of a content moderator. “Content moderator,” as a title, encompasses a wide variety of employment arrangements, and can be applicable to people who work for a diverse set of companies that may specialize in different types of content mediation/platforming. Working for subcontracting firms like Accenture, MicroSourcing, or TaskUs, and by extension for Facebook, may not be equivalent to working as an in-house content moderator for companies like Pinterest (Emerson, 2020), for instance, in the severity of their work, types of content that one may be expected to moderate, the degree of compensation one may be entitled to, or even in the availability of legal and mental health resources that help protect content moderators from the harmful exposure inherent to the work.

Subcontracting entities have their own specializations that impact the way in which they look to recruit and arrange their labour force, further subdividing the value and treatment of labour within their own firms. Some firms are regionally specific in their language provision (MicroSourcing, n.d.-c), and some tout their global reach as being a strong incentive for tech firms to hire them over others (Satariano & Isaac, 2021b). Even those tech companies that arrange in-house operations may look to outside support and supplement their in-house employees, thus participating in the broader global economic practices of outsourcing and subcontracting (Emerson, 2020).

In short, it is important to understand that "content moderator" is not a monolithic class of labour, whose labouring environment or requirements can be said to be uniform across the various regions, or whose forms of labour can be said to be distinctly identifiable. They are often enmeshed in a spectrum of labour differences, whose daily actions are informed by a number of different entities and "protocols" (Marte-Wood & Santos, 2021, p. 111) that participate in the globally networked schema of content moderation, and to better understand their plights, it is important to have a broad perspective that includes other studies of "similar low-wage, low-status fields (e.g. call centre work)" in mind (Roberts, 2014). However, within the scope of this research and with the above disclaimer in mind, I look to select testimonials and press releases made by former and current content moderators (and their legal and civil representatives) so as to stitch together a shared characteristic of the working environment and employment conditions that often impact a content moderator, while simultaneously keeping the specificity of the Manila labour construction at the forefront of my analysis. In this way, I attempt to comment on the "networked" nature of their labour arrangements, in addition to highlighting the historical connections that underlie this industry.

Another point of consideration is the limitation imposed by choosing to focus on a single platform: Facebook. In this instance, the research certainly carries the risk of providing insights that are specific to this platform only. Generalizing such insights would run the risk of eliding platform-specific nuances or flattening differences in platform governance, socio-political structures, economic models, platform affordances, and user variance; Kapoor et al. (2018, p. 543) note, for example, that most research focusing on the use of social media sites "to share and exchange information during natural disasters and critical events" made use of Twitter data for their analysis. This type of observation reflects the way in which users may approach each social

media platforms for different purposes and use-case scenarios in mind (Alhabash & Ma, 2017). While I seek to focus primarily on Facebook as the research object and content moderators working out of Manila as its subject, the collection of testimonials and documentation, even other scholarly research that I reference, all carry within them multitudes of voices from multiple global sites and various employment arrangements depending on the different BPO firms and client companies. What I learn by focusing, when I can, on Facebook then must be taken into consideration with this variance in mind.

With these concerns and limitations in mind, the specific scholarly contribution that I seek to make here is to highlight the valuable insights that can be gained by conducting a close reading of the narratives of commercialized content moderation in context with histories of building infrastructure of communication and transportation. These connections between histories of infrastructures I find relevant to our contemporary understanding of the global network that racializes to operate.

Lastly, I am wary of creating an umbrella term or reference that may, even if implicitly, lump together the histories of Chinese migrant labourers from late-1800s Guangdong to the 21st-century Filipino workers of Manila. This type of "Asian" umbrella runs the risk of replicating the very same orientalizing racist logic that I seek to analyze and argue against. The historical specificities that affected the Chinese labourers are not going to be the identical ones that affect the day-to-day operations of content moderators in Manila. However, what I do want to emphasize by paralleling each of their relationships to their labour position is the way in which the orientalizing racist logic continues to construct and operate the networking infrastructure of social media, just as they did in connecting train routes, as well as in shaping the discourse around them. For me, the most important aspect of such race logic is not only in creating a

category of "Asia" from which labour can be sourced but is in how such categorization pitches different race classes in competition with one another for the profit and benefit of the dominant society. Thus, my research specifically focuses on the historical and current tactics and discursive construction that make up the infrastructure by racializing, and gendering (Carmi, 2019; Jereza, 2021; Roh et al., 2015), certain classes of bodies. I see this type of close reading to be a crucial next step in better understanding how the history of digital networks bleeds into the contemporary mode of distributed and globalized labour flow that undervalue and make precarious human lives.

Roadmap

In this introductory chapter, I sought to introduce and contextualize the practice of commercialized content moderation, and to broaden the term coined by Sarah T. Roberts (2014, 2017, 2019) as a starting point of my research that studies how racialized bodies have been made fundamental to the makeup of the networked infrastructure of social media. To this end, I have further identified Facebook as the primary business entity and the social media platform of my inquiry, with an aim of reading the platform's business practice of outsourcing content moderation through select testimonies and journalistic coverages that detail the precarious labour conditions affecting the people who actually do the work at significant discount of their value and humanity. Throughout the rest of this thesis, I seek to highlight the discursive and practical ways in which the work of content moderation for social media platforms have come to utilize racialized bodies as a form of media filter that screens harmful contents.

In the second chapter, I conduct a deep dive into the historical accounts of the building of the transcontinental railway in the United States and explore how Chinese migrant labourers were racially interpellated into their role as an ideal labour resource and technology in building

the American infrastructure of national unity, communication, and commerce. My in-depth reading of this proto-history of networked communication calls on Armond Towns's framework of media philosophy to identify the characteristics of "Black mediality" of media, and in doing seeks to outline how the familiar colonial race logic mediatizes and technologizes the content moderators of Philippines today. By making some pointed connections between the history of the transcontinental railway and the contemporary mode of social media, I seek to illustrate how both workers (the Chinese railway workers and the Filipino moderators) were similarly racialized (1) to cheapen their and others racialized peoples' labour value, (2) to provide *timely* delivery of goods and information, and (3) to rationalize them taking on the most brutal labours of building infrastructures. Overall, by looking to the history of the American railway's use of Chinese migrant labour, I seek to build out my usage of the "network," "materiality," and "racialized media" as key terms that inform my discussion of Facebook's content moderation infrastructure and further shed light on how subcontracted labourers of Manila are being treated.

In the third chapter, I conduct a review of two of the most often cited solutions to the problem of content moderation: scaling and automation. In doing so, I seek to identify how the call to "eliminate" the human factor from the process of content moderation falls short of its imagination and, in effect, continues the historical racialization of its hidden labour force, further cementing the racial violence that undergirds the modern iteration and future imagining of the networked infrastructure. I look to highlight a difference that exists between the contemporary mediatization of the Filipino workforce and that of historical racialization of the Chinese migrants, and point to the way in which harm no longer becomes a violent consequence of their interpellation, but rather the very thing that makes the racialized workers' labour valuable to the companies that hire them overseas. The instrumental role that human labours play in the

moderation of social media platforms become all the more evident as Western tech companies look to automate their processes and fail. The evolving relationship between harm and the racialized bodies will help further underline the importance of addressing the materiality of our digital infrastructures as including vulnerable bodies that are primed to take on harm in lieu of the "contract holders" and their predominantly White, American bodies.

In the final chapter, I close this research by exploring the possibilities of resistance and trying to address how the "racialized technology," in its discourse and business practice, can be "used" to reinscribe or disrupt the racial identification that it requires (as well as the economic flow it is meant to enable). In this, I invoke Beth Coleman's call for an ethical encounter with "race as technology" as a possible conceptual framework that helps us reorient our understanding of race towards solidarity and connection, not division and violence. I provide a catalogue of various ways in which the racialized labourers of past and present have and are resisting the violent interpellation and are reclaiming the boundaries of their humanity. I then close this research by providing historical parallels that highlight not just violence and pain, but also resistance, change, and hope.

CHAPTER 2. NETWORK ECONOMY, OF HISTORY AND PRESENT

> All must and will, eventually, become devoted admirers of this gigantic enterprise, and they are now called upon for aid.
>
> -- Hartwell Carver, Proposal for a Charter to Build Railroad from Lake Michigan to the Pacific Ocean

> I am now sixty-two and I have experienced many hardships and difficulties in my life. I am proud of the fact that we Chinese contributed much to the development of transportation in Canada. Yet now the government is enforcing forty-three discriminatory immigration regulations against us. The Canadian people surely must have short memories!
>
> -- Wong Hau-hon, Chinese American Voices

Missing bodies in our discussion of "networked and digital economy"

The more conventional discussions of telecommunication networks, or the economic flows that leverage such connections and its global reach, often elide the "materiality" of its connections (Baginski & Malecki, 2013; Government of Canada, 2019; Shaw, 2006; Zwass, 2006). They instead choose to focus on the contents that circulate through the networks, letting considerations of its material infrastructures or of its connectivity enabled by human labour fade into the background, untouched and taken for granted. Nicole Starosielski makes a similar observation in her work on undersea fiber-optic networks. While transporting "almost 100 percent of Intercontinental Internet traffic" of the world, discussions of networks rarely include these undersea cables in the contemporary imaginaries of communication infrastructures:

> When communications infrastructures are represented, they are most often wireless: handheld devices, laptop computers, wireless routers, cell phone towers, "cloud" computing, and satellites pervade our field of view, directing our attention above rather than below and reinforcing a long-standing imagination of communication that moves us beyond our worldly limitations (Starosielski, 2015, p. 5).

Starosielski observed how even those that worked in the industry of designing and laying down the undersea cables expressed a similar sentiment that narrativized the extant communication infrastructure as being routes comprised of “friction-free surfaces”: “Why would you want to know? When you turn on a computer and you send an e-mail, do you really care how it works? No, you just want e-amil there, and you start drumming the table if it takes thirty seconds” (Starosielski, 2015, p. 4).

In a similar vein, many commentators on digital economics have broadly pointed fingers at various digital means and modes of engaging in market activities as comprising “digital economy”, and thus have conceptualized digital connections as extending the extant and historical market activities through digitized media, in addition to considering acts of producing and distributing “information technology (IT) goods and services, and the [generalized] effects of IT” (Government of Canada, 2019; Shaw, 2006). In this sense, the “sharing of business information, maintaining of business relationships, and conducting of business transactions by means of telecommunication networks” simply take places as an evolved form of past forms. However, this conceptualization of digital economy ignores how the circulatory flows of today take place by following “paths that are tried and true, often following contours of earlier networks, layered on top of earlier telegraph and telephone cables, power systems, lines of cultural migration, and trade routes” (Shaw, 2006; Starosielski, 2015, p. 3); and their spaces are made to be “friction-less.”

So, to centre materiality of networks in our conceptualization of social media platforms, as well as its globalized practice of outsourcing content moderation as a part of the process that constitutes said materiality, is to highlight the importance of articulating how these networks are often intentionally made “friction-less” and at a cost that is steep and disproportionate to some. It

is to emphasize that the material construction of the infrastructure of global telecommunication (and especially that of Web 2.0. and its social media networks) includes not only steel beams and glass wires, but also human bodies (Parks, 2015, p. 357). It is to further emphasize that the materiality of infrastructure does not involve all bodies, but specific bodies that are racialized to be a part of this network. It is this distinction that must take the limelight when we seek to articulate how our infrastructure is constructed.

In attempting to make the above-noted distinctions, I seek to first outline specific aspects of the condition of being "networked" and of being "material," before moving on to the work of following the "contours of earlier networks" to better understand contemporary social media network operations. In this light, I borrow Alexander Galloway's (2010) definition of "network" as being "systems of interconnectivity" that are more than an aggregation of its parts, but ones in which those parts are held in "*constant relation*" to one another, often so in a level of certain "*complexity*" (italics mine). In this definition, the quality of "complexity" helps us emphasize that our consideration of "networked, digital economy" does not simply point to its composite nature that consists of and comprehends "various parts united or connected together," but that more importantly how said parts are "connected" together through "various degrees of subordination" ("Complex, Adj.," n.d.). My emphasis is on how holding its component parts in "constant relation" can be done only by ordering each part. By focusing on these relational power dynamics that enable the modes of subordination in our networks, we can also see how and what enables said connectivity, and thus mediation (of sociality, value exchange, communication, packets of data, etc.) of the network that we seek to investigate.

As noted at the outset of this chapter, most important to my consideration of the materiality of infrastructure is the "bodies" that comprise the network. Pointing to Shannon and

Weaver, Galloway (2010) notes that “informatic networks are relatively indifferent to semantic content and interpretation.” My exploration of the various histories of digital networks that necessitated the use of racialized bodies underscores the fact that data is not simply “parsed” but is indeed read and wrote, and that a media object’s definitions, especially in relation to its permissibility of transmission, sits “at the intersection of two protocols (two technologies).” Galloway is right to point out that networks “are not simply textual entities,” and that “they are entities in a constant labor with themselves,” in as much as there is constant labour involved in maintaining the sociality of the network, and in deriving value from the network’s operation. However, this constant movement and the labour needs are not simply done by itself, or against neutral data sets that move and move through its connected protocols. Rather, the constant demand for labour is often borne by the racialized labourers who are called to process information by sensing for trauma, filtering out signal from noise, thus allowing others to connect to the broader platform that they mediate. Their labour availability is made constant, and they must be inserted as a mechanism that reads and parses specific register of data, often involving trauma, violence, and harm. In this way, the technology of race works actively to prime their body to be a preferred and even “ideal” component of the infrastructure. In this, the racialized body’s capacity to register harm (which is process that involves experiencing harm in some degree) is a core part of the technology and protocols of networked connectivity.[4]

[4] Content moderator’s critical role as a “core part of the technology and protocols of networked connectivity” does mean that someone will be inevitably called on to do the job. At the moment, taking out the human from this process (i.e., automation) is not a fully realizable possibility nor is it a desirable one. Thus the question arises, “who must do this work?” I will return to this consideration in the third chapter to explore more about automation as an alternative, and in the final chapter of the thesis to discuss “what next?” in detail.

<u>“Black mediality” in our networks</u>

It has always been a crucial aspect of the colonial and imperial project to fix its subjects to a role, a price, or a race. Subjects and their bodies were racialized, and in doing so they were made fulcrums and a measure upon which the imperial economy could turn on and off its economic activities, circulation of information as commodity, and animate its “civilized” emergence. This project in part involved the questions of who is made human (or how a human is made) in relation to the process of mediation and communication.

In “Towards Black media philosophy,” Armond Towns explores this very question by bringing together the work of Frantz Fanon and Marshall McLuhan to bear on the question of “*whose* form of mediation is ground for comprehending *whose* knowing and being?” (Towns, 2020, p. 2). In his writing, Towns explores the racial implications that previously went unexplored in McLuhan’s scholarship on media as “extension of man” by overlaying Fanon’s work on the Black body as a “vehicle toward the materialization of Western man.” In doing so, he points out how the “Western man” came into being by enacting violence against Black bodies rather than through mutual recognition of others (Towns, 2020, p. 4). Towns notes that his ultimate goal in this regard is to push “communication, cultural studies, and media philosophy toward examining the racialized and colonial implications of media technologies,” and to do so by primarily focusing on (1) providing a basis for “theorizing the connection between Western media as highly raced, contractual agreements on who/what constitutes Western man,” and (2) by bringing a greater awareness to the “wider disappearance, or forgottenness, of the Black body to Western man’s constitution” (Towns, 2020, pp. 2–3).

Towns’s intervention critically looks at McLuhan’s usage of “tribal,” “detribal,” and “retribal” framework of media and civilization, wherein the Western man was said to have come

out of their “tribal nature” of being by the virtue of his extension of self through the medium of “phonetic alphabet” (Towns, 2020, p. 1). In this, Towns notes, race was not a secondary consideration for McLuhan, even if he did not explicitly or consciously engage with it. Rather, race was central to McLuhan’s conception of time and space, as well as to his understanding of media environments. Underlying this understanding was seeing media (like orality, sound, and pictorial writing) as a mode of knowledge production, and in this context, McLuhan’s theory of media presumed “racial-techno superiority over those people of colour who did not have the same media technologies” or did not use them to the same ends (Towns, 2020, p. 6).

Pulling from J. C. Carothers, the “racist colonial psychiatrist who studied the African mind,” McLuhan established his progressivist notion of tribal (as being connected to orality and the “multisensorial reality”), detribal (as being connected to the typographic and the visually observable universe), and retribal (as being about reconnecting to the full senses of oneself through digital means) as a process of “fission of the sense” that lead to “desacralization of knowledge,” which allowed detribal man to connect to “civilization” through “typographic” medium like printed materials like books and maps: “detribal man, for whom the sense of sight was most important, was the epitome of civilization. He came out of the multisensorial tribe, in ‘linear progression,’ to become a new man” (Towns, 2020, pp. 8–9).

McLuhan, to further develop this line of thinking, pointed to Prince Modupe’s “controversial autobiography, *I Was a Savage*” in *Understanding Media*. Here, Towns remarks that McLuhan uncritically looked to the conflated figure of Modupe as someone who has come to escape one’s “tribal”-ness through mastery over the media of phoneticism, and in turn as a “*measure* of how far Western man has come out of nature, ultimately re-universalizing humanness as himself (Towns, 2020, p. 7). In developing Carothers’s racist differentiation of the

"Western man," McLuhan went on to replicate its structure, merely reading the "distinctions between Western children and African children" as relations of media, substituting for relations of race:

> […] the Western child of Carothers lived in the media environment of the printing press, the phonetic alphabet, or media that facilitated cold calculation; the African child (like Modupe) lived in a context of orality, non-phonetic writing, or relations that enabled attachment and community (Towns, 2020, p. 9).

Fanon critiques this unproblematized sourcing of Carothers's scholarship. For Fanon, Carothers's racist assumptions about the superiority/inferiority of European and African minds ignored the "colonial situation" established through violence. Instead, Carothers's analysis of the "African mind" as always playing catch-up to the Europe, and as being a "lobotomized" version of the European mind, was dependent on his own ignorance (and indeed active erasure) of the "mental and physical effect of colonialism and racism": "Fanon argued that rather than the 'colonial situation,' Carothers's study of Western phonetic literacy versus African orality made violence inherent to the Black psyche, in general" (Towns, 2020, p. 10). Then Towns, by "taking Fanon through McLuhan," underlines how McLuhan's scholarship on media is silent about the way in which the West was able to imagine itself as being distinct from nature, to see itself as a "detribal" man and a civilized man, by enacting racial and colonial violence against Black bodies. In short, their extension through media, rather than the "Gutenberg printing press or phoneticism," was through the Black bodies (Towns, 2020, p. 11).

Thus comes Towns's second critical intervention into the media philosophy of McLuhan: understanding that Black bodies are made to share an isomorphism with McLuhan's media. Black bodies, much like media, are not used for "mutual recognition between them and Western man", but rather they were rendered as "extensions of Western man." In short, Black bodies become the tool to extend Western man's sense of self as civilized – as a "detribal" being. This

recognition of self and construction of White civilization (i.e., empire) take place through racial violence against Black bodies, and this violence includes erasure of both media and the violated bodies, which are "forgotten in the making-up of Western man." This process of invisibilizing the bodies, Towns notes, is a process that imbues the bodies with "mediality," as defined by Sybille Krämer.

When talking about the "Black mediality" of the racialized bodies, it is important to understand that the usage of the concept "Black body" is not representative of "Black people, Black futurities, or Black conceptions of humanness" (Towns, 2020, p. 3). Instead, the term speaks to the way in which individuals, and indeed certain classes of individuals, can be made to "serve a function *by* and *for* others, rather than for itself" (Towns, 2020, p. 2). Towns quotes Hortense Spillers to further iterate how Black body becomes a "being for the captor" and cites Kumi Silva to talk of "blackening," a process of "'identification' of non-White bodies as open to state violence" (Towns, 2019, p. 2). Otherwise put, "Black mediality" is about the contractually agreed upon violent disappearance of the Black body behind the West's own emergence as civilization. It refers to the way in which Western man's extension of self through technology has not merely involved things like Gutenberg's press, but has also involved "race" of others which they utilized to continue "a particular genre of human (Western man) as the dominant, universalized human" (Towns, 2020, p. 20).

In reference to these textual frames, I draw upon Towns's "Black mediality" to argue that the Chinese labourers were racialized into their role as the builders of North America's first transcontinental railway. Their supposed machine-like efficiency became the "logical" reason for their employment at a historical moment when there was an ongoing labour shortage that impacted the western construction of the transcontinental railway. Eventually, using their labour

became a crucial strategy for keeping the construction costs down while adhering to the tight timeline set out by the state's contract with the Central Pacific Railroad Company (CPPR). In this historical time, the category of White and moral labour was being constructed against the foil of the Chinese male body, as differing "radically from the American male body," and thus being able to withstand more depravation and more violence inherent in the work then the "American and European laborers" (Roh et al., 2015, pp. 10–11). The Chinese labourers' bodies, to borrow Roh et al.'s words, were "a form of expandable technology" (2015, p. 11), more ideal for the dangerous work of routing American networks through its harsh terrains, and they were meant to mediate what was temporally possible at the expense of their safety. The Chinese labourers were additionally made part of the mediation process of the American unity in a post-Civil War landscape, which involved hectic (re-)balancing of the racial hierarchy in turmoil. Indeed, the process of "opening up" was not simply about racializing their bodies to be able to withstand violence, priming them for it, but it was also about making tools out of their bodies so that they may inflict violence against one another. My later argument for their place as a "racial fulcrum" come from this point of understanding of how the Chinese labourers were mediatized and made to carry "Black mediality" or, more precisely, what I refer to as Colonial mediality.

The above outline of Armond R. Towns's "Black mediality" framework provides several insights salient to my reading of the history of the transcontinental railway and, by extension, to my reading of the contemporary labour conditions affecting the globalized labour force of content moderators: (1) there is an intimate relationship between time, civilization, technology, and race; (2) there is a necessary disappearance of the racialized subject in their mediation of the "racial contract" formed between a White subject and its civilization (i.e., empire); and (3) this disappearance is often enacted through violence.

Understanding Colonial mediality

Amidst these similarities, it is important to specify certain differences that emerge in my reading of the Chinese migrant labourers' work on the railway and that of the CM workers in Manila, Philippines. While "Black mediality" provides a solid framework of understanding that (1) race as technology has been foundational to the Western man's emergence into civilization, especially in the context of American history, and that (2) such racialization has involved violent disappearance of the Black bodies, there exists notable differences in the mode of racialization and colonial violence that Towns is drawing from to formulate his thought and the history that I draw here for my thesis. While I seek to explicate such differences and nuances in detail throughout this chapter, I will provide a brief summary of the differences and specificities here.

I posit that Colonial mediality is an embracement of the imperial logic that makes certain bodies, and people, more available than others for violence and exploitation, and that this quality of media is something that racialized individuals and classes of people are made to carry as a part of their economic and infrastructural instrumentalization. Towns's formulation of "Black mediality", along with my application of Colonial mediality that mediatizes the Chinese migrant labourers and the Colonial mediality that mediatizes the Filipino CM workers, fall within this formulation of understanding how the imperial economy and infrastructure building has addressed its demand for cheap and available labour.

In the case of Chinese migrant labourers, colonial history of the Opium War made their labour available by forcing many to migrate to America. From there, they were subsumed into the labour demands of the transcontinental railway through the techno-orientalist logic of *excess*: Chinese labourers' bodies were envisioned as being robot-like, unfeeling, and able to withstand tremendous amounts of abuse inherent to the work of laying the railroads. They were thought to

be faster, stronger, and more durable than their white counterparts, while being content with cheaper wages. In this way, whether by more or less, Colonial mediality of the Chinese labourer interpellated them as the ideal railroad worker.

In the case of Filipino CM workers, this process of racialization took place in terms of equivalency or as a form of "access." The colonial history of American occupation made their labour available by transplanting many cultural and social institutions in the Philippines and preparing a workforce capable of understanding English and American cultural references. In this way, the Filipino CM workers are imagined as being similar to their American counterparts, both in their linguistic, cultural, affective, and moral capacity to experience and interpret the contents that they are tasked with moderating. The Colonial mediality of the CM has mediatized Filipino workers as another kind of robotic stand-in for their White American counterparts, this time as an avatar that can be damaged in lieu of the American users of the social media platforms who must be protected.

However, there are common qualities of Colonial mediality that affect all three formations – disappearance through violence of the racialized bodies and its people from the apparent operation of the infrastructure and from its historical records, as well as from its celebrations and their share of the successful enterprise of the infrastructure. The imperial emergence (and the logic that supported its emergence) could take place only by the dynamic application of various medialities that holds each other in "constant relation." None of the medialities are mutually exclusive in the total operation of the imperial efforts to classify, control, and enlist subordinated racial others towards its own goals, nor do they provide an exhaustive list of forms of mediality that propelled the rise of Western empires and their modern nation-state inheritors, most notably the United States.

The history of American empire is riddled with various strategies of racial management and economic exploitation that pitted different racial categories against each other, and our reading of specific histories of infrastructure must be cognizant of this cruel dynamism. The White conception of self as a civilized man came from the dynamic process of gendering and racializing others of various groups. This dynamism contributed to the "contractual value" of what was being drawn up for the imperial emergence being in a state of constant flux, especially as different racial categories were pitted against each other, at times being made to devalue each other's worth within the capitalist system of labour extraction, which also meant further entrenching classes of race in the violent system. By focusing on the specific application of "Black mediality" to the development of Colonial mediality in the Chinese labourers and to the Colonial mediality in the Filipino CM workers, I am looking towards very particular slivers of the history that collectively tell a bigger narrative about infrastructure of communication and history of American empire. I am, in part, looking to the "constant relation" that exists between these forms of medialities and strategies of power that have differentially racialized others.

Reading the history of transcontinental railway through the framework of Colonial mediality

The railway, as a networked system of communication and transportation, propelled the emergence of the American imperial economy. It has received much scholarly interest and has been the source of interrogation into how impacted people's perceptions of space and time, including ruminations on the nature of imperial economy under the auspices of emergent nation-state apparatuses and self-formation (Elmer, 2004; Innis, 1923, 2008). However, locating the racialized labour force that made it possible within its history has often been challenging because of "the absence of reliable and abundant evidence, silence, myth and lore [that] have become attached to railroad history" (Chang et al., 2019, p. 4).

By understanding the railway as one of the proto-infrastructures of communication and economic exchange, I seek to identify where the racialized bodies, and the violent process of racialization, can be found within the extant history of the transcontinental railway with an aim of extending this reading to further identify where racialized bodies are within the ongoing operations of contemporary infrastructure of social media, namely that of Facebook. Both infrastructures share a geographical point of origin: the San Francisco Bay Area. While the endeavour of building the transcontinental railway was not a singularly American project, and even though my research makes use of the Canadian scholarship around media and transportation infrastructure, my focus on the American portion of the history of the transcontinental railway acknowledges the geographical overlay that connects the infrastructure of transportation and the infrastructure of social media network. The emergence of digital corporate power out of the San Francisco Bay and its adjacent area colloquially known as "Silicon Valley" inherits much of the imperial aims of connecting the world through the exercise of colonial violence. In order to better understand this historical continuity, let us first look to how the transcontinental railway first began taking shape in the late 1800s.

With the passage of the Pacific Railway Act of 1862, the Central Pacific Railroad Company (CPPR) was chartered to build the western portion of what became known as "the first transcontinental railroad" (Chang et al., 2019, p. 1). The eastern portion of the railway, which was connected in Promontory, Utah, seven years after its chartering, ran through "vast flat expanses of prairie," making its construction a relative straightforward matter. But the western portion of the railway required the much more strenuous and potentially fatal work of navigating through the terrains of Sierra Nevada, which included tunnelling through "long stretches of solid granite" and "constructing trestles across deep canyons" (Chang et al., 2019, p. 1). The labour

force required for the construction was primarily made up of European immigrants initially, but by mid-1864, the European migrant labourers started to abandon the "back-breaking work" of railroad building and instead sought out their fortunes elsewhere, in such sectors as agriculture and mining (Chang et al., 2019, p. 2).

To fill the growing gap in their labour needs, Leland Stanford, president of CPPR, along with his fellow owners, turned to Chinese migrants from nearby mining communities. This exodus of the White European labour and influx of the Chinese labour, by late 1864, led to the Chinese workers composing the vast majority of the labour force, eventually numbering in the thousands and growing in excess of fifteen thousand over the span of the railroad's construction (Chang et al., 2019, p. 2).

The majority of these workers came from Guangdong Province in the Pearl River Delta region of southeastern China. The least prosperous regions of the province, the Siyi (four counties), produced "90 percent of the Chinese immigrants" to California in the 19th century (Dearinger, 2016, p. 153). Many of these workers were escaping an "unforgiving landscape" that impoverished entire farming communities, as well as the aftermath of Opium Wars (1839–42, 1856–60) that exacerbated already poor conditions in the region (Dearinger, 2016). Pushed by poverty and British colonial violence, these men participated in seasonal migrations in and around the city of Guangzhou, where they worked as hired hands and encountered "information networks with knowledge of and connections to work across the ocean" (Dearinger, 2016, p. 153):

> In two great waves, first in the 1850s and then in the 1860s, these rural workers followed what became a rather well-trodden route from one of the rural Siyi to Guangzhou, where they boarded flat-bottom boats to Hong Kong, signing on for passage to San Francisco. Crammed into Pacific Mail Steamship Company vessels, they endured a harrowing two-month passage to America (Dearinger, 2016, p. 154).

So, they were made to participate in a timeline that the Congressional Act, and its subsequent legislations, attached to the construction of transcontinental railway. As well as providing funding, this timeline meant that the construction required a supply of continuous and expedient labour that could meet the contractual deadline. Chinese labourers, although initially met with skepticism, came to be hailed for their expedience, stamina, and most importantly for our consideration, *speed of work* (italics mine) (Chang, 2019, p. 37; Chang et al., 2019, pp. 10–13, 18, 20).

Chinese labourers: body as a temporal measure

This imaginary of Chinese labourers as being a speedy worker, at times setting the measure and standard of work, came out of direct competition that the Chinese migrants were put up to against other migrant labourers from Europe, namely Irish and Cornish miners: "The Chinese always won. Crocker reported: 'we measured the work every Sunday morning, and the Chinamen without fail always outmeasured the Cornish miners'" (Chang et al., 2019, p. 13). These competitions were accompanied by bets, and in one instance, a commentary reports, after having won, the Chinese miners even came away "less fatigued" (Fishkin, 2019, p. 278). The discursive positioning of the Chinese labour as a measure of speed, one seemingly unhindered by the brutality of the work, continued, and in a missive to the U.S. Congress, Leland similarly discusses the Chinese labourers as figures who are "more prudent and economical [than the White labourers]," but are "contented with less wage [than their White counterparts]" (Stanford, 1865).

So, the Chinese labourer was constructed as a figure able to work faster than their White counterparts, at equal expertise but for cheaper wage, while being able to withstand the tremendous violence of the work involved in tamping granite and laying down miles of iron.

This process of racializing and idealizing the Chinese labour force helped mediate the emergence of a new spatiotemporal reality of the American empire, a new hierarchical measure of value and possibility, as much as the railway that they were constructing was meant to do. While the railway represented movement, modernity, and progress that was "deeply tied to the notions of whiteness, civilization, and manifest destiny" of America, the Chinese labour also worked "as tools of American capital" to mediate and settle the detritus of the racial anxiety that many White Americans felt at the prospect of immigration, especially at a time when the "racial status of [the White] European immigrants was far from settled" (Gow, 2019, pp. 238, 242). In the very same missive, for example, Leland notes that the Chinese labourers, with proper training "soon become as efficient as white laborers" in a line of comparative thinking and equation that illustrates how the measure of good work was always calculated on the basis of Whiteness.

Chinese labourers: body as a racial fulcrum

However, the imaginary of Chinese labourers as being the expedient and economical labour option was not so straightforwardly constructed. It required a razor-back rigour of doubly managing the rising racial anxiety of post-Civil War America and the negotiation of the positionality of Whiteness that was happening at the same time. When praising the desirability of the Chinese labour force, for example, news coverage of the time made sure to note that the employment of the Chinese "[does] not prevent the engagement of white men['s labour]" ("The Chinaman as a Railroad Builder," 1869).

In "Death of Reconstruction," Heather Cox Richardson paints a clear picture of the great anxiety that surrounded the American political discourse on labour, both in its availability and value, at the time of rising immigration from Europe and China, which coincided with the outward migration of "freedpeople" leaving the U.S. South: "Northerners associated the

newcomers with labor radicalism and worried that increasing numbers of immigrants who believed in class conflict would overawe the harmonious American free labor system" (Richardson, 2004, p. 163); "Everyone in America, North and South, black and white, was aware that a black exodus from the South endangered the Southern economy" (Richardson, 2004, p. 167). In this swelter, U.S.-born Whites, and especially its land-owning planters, were challenged to reshape their relationships with the racial hierarchy of the country (Dearinger, 2016; Richardson, 2004, p. 169).

In this nexus, the figure of Chinese migrants played a crucial role, both directly as a replacement labour for the newly "freedpeople" and as a stand-in or measure of the White migrant labour of the Irish wage workers and as a form of "leverage," a racial fulcrum, to mediate the White-Black race hierarchy now in tumult. Richardson notes that "Northern Republicans abhorred the Black Codes and other indications—like planter's attempts to import Chinese coolies to replace the free Black workers—that Southerners were clinging to the past. It was *time to move forward to a new national prosperity*, and Northerners scorned those whose primary goal was still to hurt the North" (italics mine) (Richardson, 2004, p. 29). The figure of "coolie labor," Moon Ho Jung notes, was "a conglomeration of racial imagining that emerged worldwide in the era of the slave emancipation," and represented the deep friction at the "intersections of free and slave labor, and black and white labor" (Gow, 2019, p. 240). Just as the railroad was meant to signal America's emergence out of the painful division of the Civil War and the outdated relics of slavery, racialization of the Chinese labour, whether as "coolie" or not, mediated the great anxiety that America had about its ongoing need for cheap labour, which had become the foundational logic and means of its imperial economy:

> During the Exodus, Republicans living in the East compared the suffering freedpeople with the good Chinese workers being hounded in California by gangs of disaffected thugs… Republicans in the East defended the Chinese, maintaining that they were

> model workers opposed by men hoping to extort high wages for poor work in a tight labor market (Richardson, 2004, p. 165).

When thinking about the history of the transcontinental railway, the tightrope walk that the racial economy had to perform is perhaps most apparent in the words of Leland Stanford himself, communicated in the aforementioned missive, where he stated, "no system similar to slavery, serfdom or peonage prevails among these laborers" (1865). In this context, the way in which the Chinese labourers were made "contented with less wage", while requiring them to work at an inhuman (i.e., faster than the human, measured as White) speed, makes up the logic of imperial economy and helped dictate the temporal re-arrangement that warped what was possible by the force of what was needed, and created the figure of "Chinamen as a born railroad builder." This was not simply a technical role, but also a temporal and a *measure-making* one, and fulfilling such a role came at great personal cost.

Canadian railway, information commodity

Canadian railways share much of this history, especially of the labour that was migrated over from California and China. Incorporated in 1881, the Canadian Pacific Railway company (CPR) was tasked with constructing a transcontinental railway, a fulfillment of "a promise to British Columbia upon its entry into Confederation" (Lavallé & Marshall, 2022). The railway, which was a crucial mode of connecting Canada coast to coast, was predominantly worked by Chinese labourers: "Over the course of construction and by the end of 1882, of the 9,000 railway workers, 6,500[5] were Chinese Canadians" (Ministry of International Trade, n.d.). They were employed to build the B.C. segment, a notably more dangerous terrain, and were paid 1 dollar a

[5] The Canadian Encyclopedia notes that the number totals 15,000 temporary Chinese labourers.

day, compared to the 1.50 dollar to 2.50 dollar per day that White labourers were paid (Ministry of International Trade, n.d.).

Starting with his 1923 study of the Canadian Pacific Railway system, Harold Innis penned a body of work that sought to understand and explain how the spatial arrangement of North America and its concomitant railroad networks were shaped by the imaginaries and orientation of the imperial economy. Moving through beaver pelts and cod fisheries, Innis saw the establishment of Canada as an important node being developed through a "succession of products or staples demanded by the imperial centre" (Watson, 2008, pp. xiii–xiv). Innis's scholarly focus eventually led to his examination of pulp and paper products as the mediating staples of the exploitative relationship that the imperial centre (located in Europe, especially through metropoles like London and New York) had with its peripheral nodes (i.e., American frontiers and regional townships) as they impacted materially the "demand caused by the advent of the mass circulation daily newspaper" (Watson, 2008, p. xvii). The rising speed with which news could circulate, thanks to ready resources of paper, was both necessitated and propelled forward by the increasing centralization of information that shaped the operational blocks of the imperial economy like its systems of markets, price information, and credits (Hever & Crowley, 2008, p. xxxi). Hever and Crowley (2008, p. xxxix) talk about how the emergence of such facets of economic exchange led to the information itself taking on the role of commodity by saying "monopolies of knowledge created another kind of dependency relationship between the centre and the margins." These were Innis's observation of what he referred to as "space-biased" civilization which created hierarchical social orders that allowed elite groups to exercise "exclusive access to a monopoly of knowledge…which in turn became a powerful tool in regulating the division of labour of the populace" (Hever & Crowley, 2008, p. xxxiv). The

expansionist empire subjugated marginal groups and through violence administrated over great distances, in part with the help of advances in telecommunication technologies, but also by moving those bodies across oceans and through the harsh environments of the North American plains and mountains, so that they could lead to the "emergence of complex political authority and [contribute] to the creation of abstract science and technical knowledge" (Hever & Crowley, 2008, p. xxxv). For Innis, this triad of economy, time, and technology was at the base of any empire.

However, what is missing for us, similar to what Towns found lacking in McLuhan's consideration of media, is the acknowledgement that the imperial biases were formulated over great spans of distances and in between betwixt times by fixing racialized bodies to a function of the network. They were made to serve as a form of media that produced knowledge about the economy, race, and various relationships that undergirded the rise of empires. Not only were the Chinese labourers tasked with blasting the hard granite of Western Sierra Nevada with explosives and felling trees that stood in the way of the railway, thus helping Western men control and quell the natural landscape towards the goal of connecting a "civilized" nation, but the labourers also helped mediate the colonial relationship between White settlers and Indigenous peoples. Shinya Huang, in his reading of the Native American tribal archives and interpretation of oral history relating to the accounts of Native American encounters with Chinese labourers, notes how the construction of the railroad led to the forced removal of Native Americans from their lands (Hsinya, 2019, p. 180). Moreover, Hsinya notes how the "incomprehensibility of the Chinese labourers names" became one of the justifying rationales for structuring their payment schemes (i.e., not paying them individual wages), a tactic later extended to the Native Americans employed alongside the Chinese labourers, thus impacting

their wage structure as well (Hsinya, 2019, p. 182). This mode of racial management was utilized alongside the railroad companies' exercise of the federal land grants that "eradicated Native American claims to land along the tracks" (Hsinya, 2019, p. 186). In exercising their contractual rights and obligations with the state, the railroad companies mediatized the Chinese labourers' bodies to produce knowledge about the environment they were boring through, as well as the Native American people who were being displaced and simultaneously exploited as a source cheap labour. In places like Donner Summit Camp, one of the earliest examples of a Chinese labour camp in the United States that ran along the railway during its construction, the landscapes of the North American West that were once considered "hostile" and inaccessible became better known to "settlers and newcomers to the region" through the presence of workers and their shelters (Dixon, 2019). Here, we see the Colonial mediality of the Chinese labourers' bodies clearly, in how they mediated the racial contract that gave rise to the emergence of the White self that Towns notes (Towns, 2020).

Chinese labourers: erasure of the bodies

If "figuring" Chinese labourers into such a schema of racial management was a way of materializing the railway infrastructure, then so was their erasure from the histories of the transcontinental railway. Chang and Fishkin note how challenging it was to conduct their research in the absence of historical documentation on the Chinese migrant labour force, including the lack of records left by the labourers themselves: "Business records, including those from the Central Pacific archives, are incomplete, scattered in different locations, disorganized, and difficult to decipher" (2019, p. 4). Neither this disappearance nor the absence of textual records, Chang argues, is for the want of literacy among the labourers. Rather, such absence denotes the "colonial condition" and the racist violence against the Chinese homeland and the

diasporic communities in the United States that bear the legacy of this silence so many decades later (Chang et al., 2019, p. 4).

This big historical lacuna then is characteristic of the condition of "mediality," which Towns borrows from Sybille Krämer's work describing how media disappears in the very act of delivering messages: "We do not hear vibrations in the atmosphere but rather the sound of a bell; we do not read letters but rather a story" (Enns, 2015, p. 11). Anthony Enns's introduction to Sybille Krämer's work, "Medium, Messenger, Transmission," quotes Krämer's earlier work and highlights this specific quality of medium clearly: the medium is neither seen nor heard if it is properly functioning in its role as a medium. A medium only becomes apparent in its failure. Thus, the correct operation of the colonial power not only called on the Chinese migrant labourers to work the violent work of constructing the railway, it also disappeared the workers behind their work.

Upon the railway's completion, Leland Stanford invoked the "Siamese twins" as a metaphor that described the way in which the railway "formed the 'ligament' that binds the Eastern Eng and Western Chang together" thus emphasizing the international appeal of the railway (Fishkin, 2019, p. 351). But simultaneously, Andrew Russell's iconic photo, "East and West Shaking Hands at Laying Last Rail," depicts Leland driving the "last spike" but none of the Chinese labourers who laid down the final yard of the "iron road" are visible (Chang et al., 2019, p. 17). Soon after, the newspapers reported that the shipment of tea from China started eastward at the same time as the last spike was being driven in (Chang et al., 2019, pp. 19–20), but Judge Nathaniel Bennet, a key speaker at the celebration in San Francisco, hailed the "Californians" for having built the railroad, thus couching the railroad project in the language of nationality and citizenship, from which the Chinese labourers were denied (Chang et al., 2019, p. 33).

This simultaneous presence and disappearance are emblematic of the Colonial mediality of Chinese labourers and contribute to the difficulty of accurately and fully comprehending the material make-up of networked infrastructure as including racialized bodies. Their presence was tied to the railroad's construction and its capacity to move consumer goods, thus animating the American imperial and transnational economy through a "racial contract"; but such movement only made manifest in the service of the empire's conception of itself as a united, White state. The extreme toll of such mediation is perhaps most bleakly exemplified in the fact that one of the most lucrative exports from the United States in the mid-to-late-nineteenth century was the "cleaned and prepared physical remains of Chinese who had died in the United States and were returned to China for final burial" (Chang et al., 2019, p. 40). The very same colonial network that brought Chinese migrants into America was used to export out the "remains" of the violent process that rendered Chinese in utter negation and abject "disappearance," extracting value to the very bitter end.

Reading commercialized content moderation through the framework of Colonial mediality

The social media infrastructures of today, and the digital networks they form, inherit this colonial logic of race, as they seek to provide digital connectivity to a global userbase spread out over vast distances, with constant platform availability, and all the while keeping their operational costs down. As users flood the platforms with contents, the social media companies compete to provide not only the most valuable digital service but also a media environment that is free of harmful contents that may damage the user experience. The burden of "cleaning up" the media environments of social media platforms have fallen on the shoulders of the most underpaid and undervalued workers in the social media industry, and the Western tech

companies and their subcontracting BPO firms have collectively looked to parts of the world that can be made to meet the incredible demand for speedy, constant, and cheap labour.

In the Philippines, they found their "ideal workers" that could be made to embody the Colonial mediality required of their infrastructural design and through suitability that has been violently shaped in the likeness of the west by the history of American colonial domination of the region. Just as the Chinese labourers were employed to mediate the need for speedy construction and to bring down labour costs in competition with other racial groups in nineteenth-century America, Filipino workers employed by BPOs are also made to embody the Colonial mediality that makes their labour available 24/7 for the benefit of the Western tech companies that seek to have its services and platform available to the global audience around the clock. At the click of a button and at an incredibly attractive pricing, the globalized process of outsourcing has made their labour available for digital "cleansing." Just as the capital interests of CPPR employed the Chinese labourers to take on its most dangerous work, contemporary social media companies like Facebook employ its global labour force to trawl through its most harmful contents online. While the work that content moderators do differs from the work of the railroad workers, their labours are structured and rationalized through similar logic and means that share an "isomorphism of McLuhan's media" (Towns, 2020). Understanding these shared characteristics would help further outline how the materiality of digital networks have continued to make racialized bodies foundational to its makeup, just as the infrastructure of railroads have done. Below, I look at how those characteristics emerge in the context of the content moderation industry and how the Filipino workers are "made suitable" for the work as content moderators.

As noted in the first chapter, content moderators are often sourced based on their capacity to read and understand a platform's contents, which is a prerequisite for accurately determining

and explaining the permissibility of content. This requires a well-rounded knowledge of languages, cultural contexts, and platform affordances, as well as the capacity to synthesize multiple knowledge bases into a decision-making process, all the while undertaking this work from locations remote from where the content originates or is destined to be viewed (Roberts, 2019, p. 17). While the overall guidelines of what counts as permitted speech (e.g., community guidelines) and what counts as censorable content comes from the platforms themselves, the deciphering and application of said guidelines ultimately rests with content moderators themselves.

In this, an ideal content moderator is envisioned as a linguistically competent and culturally literate individual who in some sense shares the "mental trajectory" of the users that post the contents and the platform programmers that draft community guidelines (Manovich, 2001, p. 74). Lev Manovich uses the term "mental trajectory" to talk about a "new kind of identification appropriate for the information age of cognitive labor" in his book *The Language of Media* (Manovich, 2001, p. 72). Manovich notes that "interactive media asks us to identify with somebody's else mental structure," and that a "computer user is asked to follow the mental trajectory of a new media designer" (Manovich, 2001, p. 72). In a similar way, content moderators are called to anticipate, understand, and presume the design intent and thought processes of those that drafted the policies and community guidelines for the platforms that they moderate. This requires a complex cognitive labour that goes beyond simply reading and understanding the letter of the policies that they receive. What is most notable (and concerning) in this is the fact that the "mental trajectory" of the policy writers are imagined to be accessible by the content moderators, because the moderators' lives, histories, and education have all been shaped by the colonial arrangement of space and economy and through a practice predicated on

what Roberts describes as "long-standing relationships of Western cultural, military, and economic domination" (2019, p. 18). Social media platforms and their subcontracting entities exploit places like Philippines as rich sources of cheap and *suitable* labour, precisely because the history of American colonial dominance has fostered, by force, "a proximity to American culture" and the general "Filipino/American relatability" through a "benevolent" educational system that has become a foundational aspect of the "value-added aspect of the Filipino BPO" for US-based tech companies (Marte-Wood & Santos, 2021, p. 105; Roberts, 2019). Not only are Filipino workers constructed, both historically and discursively, as a workforce that can understand English (among other language requirements of the platform) and decipher cultural references that predominantly emanate from the Western culture sphere, but they are configured as workers who share an affective, cultural, and linguistic proximity with the majority of its user base. If the Chinese railroad workers were cast as an ideal railroad worker because they, unlike their White counterparts, did not feel pain and could work on and on, robot-like, then the draw of Filipino workers for the Western tech companies comes from the fact that they are familiar, so much so that Filipino moderators can know and pre-emptively filter out content that would have otherwise harmed their American counterparts and the average user.

TaskUs: temporal and spatial arrangement of the world

Some subcontracting entities have made this kind of "isomorphism" an important marketing point of their services. TaskUs is an American BPO service firm that provides content moderation service for companies like Facebook. TaskUs boasts of having over 27,000 employees from over eight countries and 18 locations around the world, but more than half of those locations are found in India or the Philippines, with approximately 19,190 employees, which represent about 70 percent of the total employee count, based in the Philippines (Chugg,

2021). As of the filing of its SEC IPO documents and amendments in June 2021, Facebook is TaskUs's biggest client, generating more than 30 percent of their revenue in 2019 and 2020 (Chugg, 2021). In this way, through the services of the subcontracting BPOs, platforms like Facebook can source their necessary labour force at a legal arm's length, protecting their own people while operationalizing and scaling certain functions of their platforms.

A significant part of TaskUs's success and its subsequent capacity to financialize through Initial Public Offering (IPO), they admit, is thanks to their heavy reliance on "international operations, particularly in the Philippines and India" (Chugg, 2021). TaskUs cites the generous policies of the Philippines and Indian governments to promote foreign investment as key its success in those countries. This is cited specifically in relation to the labour cost in other parts of the world ("United States and other developed countries"), which allows TaskUs to "competitively price our solutions" (Chugg, 2021), taking advantage of tax incentives and concessional rates established by the Philippines Economic Zone Authority (PEZA). Even as TaskUs seeks to expand its operational bases to match the increasing call from clients to expand their "global presence and multilingual capacities," TaskUs notes that their expansion will primarily focus on the established international sites of the Philippines and India, expanding into their respective "provincial cities" (Chugg, 2021). This presents a curious expansion of the global growth that is overlayed atop growth of the Philippines' economic forces domestically.

The underlying logic of this "competitive" pricing, however, is perhaps most clearly revealed by one of TaskUs's cited concern for future risks:

> The trend toward near-shore and offshore outsourcing, international expansion by foreign and domestic competitors and continued technological changes may result in new and different competitors entering our markets. These competitors may include entrants in geographical locations with lower costs than those in which we operate. (Chugg, 2021).

Roberts's interviews with Filipino CCM workers note how their increased productivity metric and decreased labour pool have been due to the rising competition from "other vendors" of Indian origin (Roberts, 2019, p. 180). Just as the Chinese labourers were pitted against other European migrant labourers or the Black labour force available in the post-Civil War era, TaskUs and other outsourcing firms are leveraging the globalized economic network to pit various locations around the world to eke out the most *value*, while continuing to expand and invest in regions of the world where they can benefit from a comparatively cheap workforce. In this sense we see the Colonial mediality of the Filipino labour force at play, in mediating the value relationship of the global network supply of labourers. They are not only made to participate in the global communication network of exchanging messages and audio-visual contents as its moderators, but their cheap labour is enacted alongside other valuation mechanisms like the currency-exchange rates to mediate the overall operation of outsourcing and its profitability: "Changes in other income are driven by our exposure to foreign currency exchange risk resulting from our operations in foreign geographies, primarily the Philippines, offset by economic hedges using foreign currency exchange rate forward contracts" (Chugg, 2021). The value of their labour seems to partly come from the way in which they can be undervalued more so than other racialized bodies, a process that runs parallel to the process of the mediatization of Filipino workers.

MicroSourcing: managing time and space

Another example comes from MicroSourcing, a BPO service firm also operating out of Manila, founded in 2006 by a Dutch businessman Philip Kooijman. After going through several acquisition processes, in February 2020, MicroSourcing was acquired by Probe Group, an Australia and New Zealand-based "customer experience" company (MicroSourcing, n.d.-a). In

its promotional material, the company advertises itself as an efficient, cost effective, and scalable offshoring solution for companies worldwide. But most importantly for our purposes is its description of why such outsourcing should target the Philippines (MicroSourcing, n.d.-c), a country "with just the right combination of qualities." These qualities include language competency ("The Philippines is the third-largest English-speaking country in the world" and "Filipinos don't just speak English; they speak it very well"), low cost of living leading to low operational overhead for the client companies (which are explicitly noted as being from either Australia or the US), and the availability of "huge talent pool" of highly educated workforce ("with more than 41 million highly trained people in the workforce, the talent pool in the Philippines is deep and very rich") (MicroSourcing, n.d.-c). But perhaps the most telling part of its self-promotion is that it explicitly points to the Philippines's history as a US colony as a priming factor for the "culture fit" required of a workforce able to read and review Western contents:

> As a former US colony with a 90% Christian population, the Philippines has a very "westernized" culture. Its cities, office buildings, malls and restaurants all feel very familiar to most Americans and Australians. And because there's no language barrier or real workplace differences, working with Filipinos feels almost identical to working with Americans or Australians (MicroSourcing, n.d.-c).

There is an increasing awareness and call for having regionally specific and sourced content moderators, however these are practices that seek to replicate the divided worldview along the borderlines of socially produced spaces. MicroSourcing, a multi-service business process outsourcing firm based in the Philippines, touts its employees' excellent command of English and their immersion in Western popular culture as a selling point for their content moderation services. Their website asserts that Filipino people possess innate qualities making them particularly adept at supplying moderation services for North American clientele, including a "great eye for detail." Here, outsourcing is not just a geographic concept denoted by physical-

world spatial characteristics or constraints, but instead outsourcing denotes a set of labour processes and configurations, and a designation of an available labour pool, as opposed to simply location, which is in part responding to the historical forces that have shaped labour availability. It is a type of work, a wage level, and a class of worker whose peripheralization is enabled and enhanced by digitization and the internet's "space of flows" that continues to "move" people through "well-worn circuits established during periods of formal colonial domination and continuing now, via mechanisms and processes that reify those circuits through economic, rather than political or military, means" (Roberts, 2014, p. 196). Just as the Chinese labourers of the late 1800s were moved from the shores of Guangdong to California and Vancouver, the global workforce are made to circulate and in turn enable circulation of digital goods, value, and labour. Even if they do not have to move from their offices in Manila, or even from their home computers, the racialization interpellates them over great spans of distance to energize the Western circuit.

The above closing of spaces is part of a maneuvering that affects the spatio-temporal arrangement of the world through racialized bodies. Just as the Chinese labourers were used as a time-measure for mediating a speedy and expedient delivery of the railway's construction, the content moderators are often working against a very strict arrangement of shifts and censorship quotas tailored to the needs of the platform's user base. Many moderators have reported working under extreme duress to screen content (1) at a fast pace, and (2) with great accuracy (O'Connell, 2019): "'I like to take my time, and get it right,' Plunkett says. But if a Facebook employee disagrees with her decision, or she takes a second longer than her allocated break at any given month, she says she will slide down the tier scale of moderators from Tier A to B or C and be paid less" (Nast, 2021). The moderators are subjected to a punishing evaluative framework that

demands an increasingly high level of accuracy of filtering, which is set by contracting companies like Facebook (Silver, 2018).

This extreme pressure for response time is tied to the fact that Facebook is a 24/7 international service: "If someone reports a Tagalog-language post in the middle of the night in the Philippines, for instance, there will always be a Tagalog-speaking reviewer — either locally or based in another time zone — that the report can be routed to for quick review" (Silver, 2018). Such "uptime" requirement has connected the availability of the platform to the availability of moderators with specific language proficiencies as well as cultural understandings necessary to the analytical work that undergirds the smooth operation of the platform and ferrying of contents. In this way, the racialized labour force is called to mediate the constant availability of the platform's services, similar to the way in which Chinese labourers were made to mediate the temporal measure of the American nation-state and its need for a transcontinental railway. Their mediation also underlies some of the legal and contractual agreements that the companies have entered into with various regional governments; in 2007, Facebook settled a child safety probe in New York with the New York State Attorney to ensure that pornographic posts flagged by users are taken down within 24 hours (Reuters Staff, 2007; Satariano & Isaac, 2021a).

This incredible demand for the labour's availability is tied to the company's capacity and willingness to arrange and organize labour at the expense of the safety and comfort of the employees. TaskUs, for example, notes in its 2021 Securities and Exchange Commission filing that in the midst of a global pandemic, the company, at the request of its client companies, started to bring its workforce back into the office after work-from-home period (Chugg, 2021). The mobility of the labouring bodies and its health seem to be entirely left up to the demands of

client companies like Facebook, who seek to have their platforms operational around the clock, no matter what.

Disappearance and paradoxical representation

Another core aspect of "Black mediality" (and Colonial mediality) is how media, and in my consideration the infrastructure of communication by extension, disappears in its transmission of the message. One of the core techniques of silence is through the usage of the non-disclosure agreement (NDA). Many firms enforce these legal documents that prevent employees from speaking out and seeking a public avenue for support, going so far as to prevent them from discussing the terms of their labour and the daily realities of their work with their loved ones or close contacts (Dwoskin, 2019; Gilbert, 2019, 2021). At times, these NDAs are invoked against former content moderators that have come out publicly against the horrendous labour conditions imposed by their former employees, so as to pressure or outright silence them (Nast, 2021).

In her research, Roberts (2019) makes note of the difficulty that NDAs imposed on her efforts to learn more about the ongoing practices of content moderation and the significant personal and social harm that the content moderators risked in deciding to be interviewed. This type of "friction" is reminiscent of the type of barriers to research that Chang and Fishkin (2019) faced when conducting their research into the lives of Chinese labourers on the railroad. The scant historical documentation belies the active silencing or erasure of the workers on and off the labour sites, and the looming threat of socio-legal violence that work to sanitize the history of the racialized presences foundational to the makeup of communication networks.

The use of NDAs is part of a broader complex of business practices that invisibilize the infrastructure of content moderation for social media companies. Various outsourcing firms have

engaged in practices of misrepresenting the labour conditions and portfolios to potential employees, thus recruiting content moderators under false pretenses (Block & Riesewieck, 2018; Perrigo, 2022a). They may use umbrella terms like “call centre agent” or omit the possibility of “regularly viewing disturbing content that could lead to mental health problems” (Perrigo, 2022a), thus hiring employees with incomplete disclosure. Some have even reported being asked to sign a non-disclosure agreement prior to being informed of the details of the actual work that they would be doing daily, and feeling that it was “too late to turn back,” since they had already accepted the job and had moved countries for the work (Perrigo, 2022a).

This type of lack of disclosure and intentional obfuscation of labour conditions during the recruitment process adds another dimension to our consideration of the complexity of digital infrastructures. As well as providing increasing barriers to research into the workings of the digital economy, these kinds of business strategies cloud our general understanding of how globalized labour practices continue to “move” people for profit. Digital humanities research today must be able to contend with this kind of “black-boxing,” and be able to synthesize the technical obscurity that prevents outside-in looking at the digital algorithm and platform designs alongside the matters of digital “invention” as being an aggregation of human decisions informed by social, historical, and economic motivations that surround them as well (Brown, 2010).

As we learn from the history of the railway, this kind of disappearance is coupled with the paradoxical presence of discursive representation that simultaneously hides and forgets racialized bodies. Steiger et al. (2021) note how at times CCM has been portrayed as “stigmatized work of last resort, a ‘dirty job,’” but that such a portrayal is in tension with the nobility ascribed to work that “keep[s] the internet safe from others, which workers can take pride in.” Just as racialized labour was necessary for the symbolic construction of the networked

infrastructure as delivering more than the sum of its contents or information or transport, in mediating the grander meaning of unity, connection, and human evolution (i.e., civilization), the moderators and the railway labourers seem to be made to disappear behind what they signal – that is, behind the metaphor and the cypher of meaning, and behind the infrastructure that they comprise and animate.

Over the course of my research, I have come across many interviews and statements made by the former and current content moderators or by their supporters who have expressed similar sentiments and self-perspectives (Block & Riesewieck, 2018; Gilbert, 2019; Gray, 2019; Nast, 2021; *Open Letter from Content Moderators*, 2020). Some have described themselves as the internet's "policeman" whose main goal is "to make the platform as healthy as possible." Some saw themselves as "security," protecting the users; others described themselves as digital cleaners: "the main goal of the content moderator is to clean up the dirt" (Block & Riesewieck, 2018). Some have noted the gravity of their work and the outsized impact that they can have in their position as a content moderator: "As content moderator you cannot make one mistake, it could change peoples' lives for the negative, trigger war. Our job is not easy" (Block & Riesewieck, 2018). While prominent, these are not the only modes of self-description that content moderators and their supporters have used when contextualizing and explaining their work's importance with respect to the media environment of social media networks. Marte-Wood and Santos (2021), for example, open their research on content moderation with their own close reading of the documentary *The Cleaners* (2018). In it, they comment on the film's use of religious imagery and one moderator's use of religious language in their self-description: "I'm a preventer. So you need to prevent sin from being out there on social media." Social media companies or the BPO certainly would not describe their work or the platform's functionality in

those terms. This scene illustrates the "palimpsestic layers" that form not only "overdetermining the Philippines as a site for this uniquely twenty-first-century labor," but one that also points out the complex navigation of self-description that its participants undertake in narrativizing themselves in relation to the harmful work (Marte-Wood & Santos, 2021, p. 102).

While it is important to understand how these modes and language of self-perceptions contribute to the overall industry and apparatus of content moderation, for my research, I am particularly interested to observe how content moderators (and their work) are externally described by journalists, tech companies, and industry professionals as being "essential" or as "unrecognized heroes" and, most importantly, as "a resilient workforce." Many, including the companies that hire the employees, have described the work of content moderation as being the "first line of defense" in the realm of social media, and thus playing an indispensable, albeit rarely discussed, role, all the while refusing to make more details about their labour conditions public under the guise of protecting them (Guevara, 2021; Matsakis & Martineau, 2020; Silver, 2018).

Trauma and narratives of resilience

One of the ultimate disappearances of the racialized bodies in the context of social media content moderation, then, is in the way in which the work of CCM has been a source of lasting, and at times final, harm to those that labour in the role. As mentioned previously, content moderators who have had prolonged exposure to harmful contents have reported experiencing various degrees of post-traumatic stress disorder (PTSD) from the "secondary trauma and eventual burnout" (Gilbert, 2019; Steiger et al., 2021). In some extreme cases, there have been reports of heart failures or self-inflicted harm that has lead to the death of the moderators (Block & Riesewieck, 2018; Newton, 2019). Considering these costly impacts, how should we think

about content moderators' role as the internet's "cleaners" and "defenders," and by extension to justify, rationalize, or to even demand that these roles continue to exist to enable the continued operation of social media platforms?

I argue that this type of language that describes the necessity of the labour has also come to racialize the labour and naturalize the labourers' position in the communication network and the harm inherent in the work, thus not only capturing them in a certain labour position, but prolonging their stay in it despite the harm it causes. One of the most telling aspects of this racialization is how one's presumed ability to "withstand" trauma has become a core part of who is "suited" for the job of content moderation. Then vice-president of Operations at Facebook Ellen Silver's 2018 post makes this explicit by noting, "This job is not for everyone – so to set people up for success, it's important that we hire people who will be able to handle the inevitable challenges that the role presents. Just as we look for language proficiency and cultural competency, we also screen for resiliency" (Silver, 2018).

Resiliency, and the related concepts of "mental health" or "well-being," has become a prevalent way of discussing the maintenance and cultivation of an ideal content moderation team. From business publications (Chugg, 2021; Meta, 2021a) to academic research (Santa Clara Law, n.d.; Steiger, 2020; Steiger et al., 2021) to BPO hiring posts (MicroSourcing, 2021), resiliency is mentioned as being one of the crucial aspects of what makes a good content moderator. However, this language is not so neutral, nor is it merely descriptive of an existing condition within prospective workers. Rather than referring to an individual's capacity to withstand, manage, or otherwise receive trauma, the notion of resilience must be understood as a required quality applying to broader labour pools, which functions alongside other requirements, namely language proficiency, shared cultural context, and willingness (whether by coercion or

not) to take on the labour. The groups that do the work are in some sense retroactively labelled as "resilient," and those who burn out are considered to be lacking in the quality.

Many former content moderators have pointed out that despite companies' various assertions that they are investing in mental health and wellness resources, including hiring a team of dedicated on-site counsellors, what is actually available to the moderators has been inadequate or contrary to the PR messaging by the companies. But even if we are to accept at face value the assertions of the companies that the mental health support is truly available and plentiful, the resiliency framework simply ignores the "colonial situation" that has made such mental health support necessary in the first place. What are we to make of the statistical reality that the majority of the global supply of content moderation comes from the Philippines taken alongside companies' claims that they only hire people who can handle being traumatized? It seems evident that labour of content moderation is organized so that certain class of racialized bodies must be involved, and this design of the platforms is completely ignored when we simply focus on making the labouring bodies more resilient to harm, instead of considering why they must be the ones to be put in the path of such harm. Just as the Chinese labourers' bodies were considered to be able to withstand the tremendous abuse and violence that went into flattening the granite mountains with explosives, the global labour force of content moderators, many of whom come from historically marginalized (and colonially impacted) regions of the world, are now made to withstand the psychological harm on an incredible scale, as if they alone possessed such "capacity."

Even academic works that find the myth of progress involved in creating the American ICT infrastructures problematic have at times resorted to the celebratory language that describes the workers and their bloody labour as an enterprise of the resilient: "These immigrant workers,

though that did not look, speak, eat, or spend their leisure time in ways recognizable to their ostensible superiors, had earned respect as an army of courageous, industrious, and *resilient* men" (italics mine) (Dearinger, 2016, p. 175). It is tempting to want to find ways to celebrate those exploited for their labour. Such framing instead perpetuates the discursive positioning of their labour as being mediated through their capacity to accept, register, and *endure* the violence inherent in the Colonial mediality that they were made to carry (Slack & Wise, 2015, p. 27).

Addressing the harm of content moderation simply as a problem to be solved or ameliorated (i.e., a wellness to be managed) is to accept reduction of public life to "problem solving and program implementation, a casting that brackets or eliminates politics, conflict, and deliberation about common values or ends" (W. Chun, 2018, p. 74). Wendy Chun's observation here, in part, came from noting how the individualized scores of the network users "coincide with 'older' racial and class categories" (W. Chun, 2018, p. 75). However, I think the insight can be extended to our consideration of the very infrastructure that creates those scores. Just as user profiles indicate far more than individual actions and are created by "comparing and lumping together" with other user actions through network categorization (W. Chun, 2018, p. 75), creation of a 'resilient workforce' look at various markers and "systems" that come together to create a "category" of bodies and people that are said to be 'suited' for the work of content moderation (Nuttman-Shwartz & Green, 2021). In this way, resiliency cannot be neutrally read as a characteristic that individual workers are said to possess. Rather we are called to consider resilience as a complex of knowledge about racialized bodies as an extension of colonial logics of value and message circulation, with accompanying strategies of labour that are particularly cruel to select classes of people. In this observation, I echo Rae Jereza's caution that consideration of content moderators' work as a form of affective labour should not presume that

such labour is more important than user's activities (Terranova, 2000). Rather, my observation seeks to expand on Jereza's understanding of affective labour to point out that there exists an intimate relationship between "monetisation of users' relationships, interactions and creativity" and the Colonial mediality that undergirds such interchanges on social media platforms (Jereza, 2021, p. 930). In certain ways, we do not get to message our loved ones in part thanks to the resilient bodies. Instead, we have created these bodies in order to communicate through these media environments, through the digital channels that span vast distance and complex time.

Conclusion

Through the bodies of its racialized labourers, both networks of railroad and social media have sought to exercise profitable control over the processes of production (of knowledge and of the infrastructure itself) and of the transmission of knowledge and goods. These measures of control, often inherently violent, have dehumanized the people interpellated into its process, and as a part of its operational protocol have mediatized them as the very network that manages the "constant relation" of its connections. At times, these networks have come together to span vast distances, and at times they worked to compress time and provide near-instantaneous connections or value transfer around the clock. But they always eked out value through putting racialized bodies in harm's way. Despite the two networks' historical and technological differences, the need for racialized labour and the violent imposition of Colonial mediality has remained a constant that formed the infrastructural basis of the American nation-state's imperial emergence and of the financial success of the American corporations that have enjoyed an intimate relationship to the nation-state understood as a connected power.

As a part of these visions of progress, instrumentalized bodies were made to disappear behind the apparatus of networks that they constructed and animated, being made to fully

embrace the core aspect of their Colonial mediality. In the case of the Chinese railroad workers, their bloody contribution to the railway infrastructure were sanitized out of the historical accounting of the construction and their presence post-construction was even criminalized or heavily taxed. In one of the very few remaining records of Chinese railroad workers, for example, Wong Hau-Hon reflects proudly on his work on Canadian Pacific Railway but laments the tightening of Canadian immigration law that made Chinese immigration illegal (Chan & McIntosh, 2022; *Chinese Exclusion Act (1882)*, 2021; Yung, 2006). The politico-legal instruments like the American Chinese Exclusion Act of 1882 or the Chinese head tax levied under The Chinese Immigration Act of 1885 in Canada further enabled these two nation-states to erase the humanity of those that made their networks possible.

This type of violent disappearance and negation of racialized labourers' contribution continues today and has become one of the central conditions of today's globally distributed digital networks. Facebook and other companies that provide similar media environments are starting to publicly acknowledge, albeit in limited scope, the problem of maintaining the current type of labour arrangement to support their content moderation practices. This issue has provoked questions concerning the long-term sustainability and viable growth trajectory of platforms like Facebook, and in an effort to provide a solution to the problem, Facebook has often proffered machine-learning as a future alternative to large-scale employment of human content moderators. In the next chapter, I seek to explore Facebook's big-picture solutions regarding scaling and automating, and outline the discursive and practical reality of the vision that purports to lessen human cost of content moderation. In doing so, I argue that, contrary to Facebook's self-portrayal, Colonial mediality continues to form the basis of its future vision and that the solution of automation brings in more racialized bodies into its labour arrangement while

hiding their presence in the technological make-up of its solution, thus replicating the "various degrees of subordination" in noble but ultimately discriminatory ways.

CHAPTER 3. AI-SOLUTIONISM AND RACIALIZED HUMAN-MACHINE PARTS

> A robot may not injure a human being, or, through inaction, allow a human being to come to harm.
>
> -- Isaac Asimov, I, Robot

> AI will not save us.
>
> -- an unnamed Facebook employee, internal memo

New bodies in our discussion of "networked and digital economy"

Previous chapters have outlined how the rationalized relationship between racialized bodies and the violence that accompany the construction of communication infrastructures has continued to embed colonial histories into the make-up of the media environments of various networked infrastructures since the days of the transcontinental railway. The contract holders that benefited from having such infrastructures built (i.e., CPPR and Facebook/TaskUs) sought to meet incredibly demanding labour needs by integrating racialized, expedient classes of people. The work of constructing and operationalizing the infrastructure of communication was and is incredibly demanding, often exacting physical toll upon the bodies and lives of the workers. The racist logic that undergirds such interpellation have discounted the labour and humanity of those so racialized, and in this process various classes of people were pitted against each other and thus were made to help reinforce and reiterate racial hierarchies in America. This process was ultimately made possible by putting the most vulnerable in the most dangerous situations under the racist pretext that they can handle the abuse more so than their White counterparts. This process of tamping the iron road and wiring the digital connections of 21st century, also often involved erasing their presence from the very networks that they helped create, leaving a lacuna

in the historical accounting of the railway and strategically silencing those vulnerable workers, thus preventing from seeking help and solidarity.

While all these elements of Colonial mediality form the basis of these workers' (both historical and present) brutal racialization and dehumanization, today's social media platforms call on its racialized workforce to input their bodily presence into their digital infrastructure in a slightly different way that not only makes harm done to their bodies a haphazard consequence of the interpellation, but as the very reason for it. From cultural familiarity (which harnesses the colonial history of American domination of the Philippines) to resiliency, the Filipino content moderators are assigned their role as an "ideal worker" through various discursive configurations that make their bodies available for the work of sensing harm and thus being (and otherwise having the capacity to be) traumatized. In today's world of globalized labour, they are becoming all the more foundational to content moderation processes and social media's platform design with the help of subcontracting businesses (i.e., BPOs) that partner with Western tech companies in their search for a cheap and available workforce.

This chapter explores two major forms of artificial intelligence (AI) solutionism that have come to form the basis of contemporary discussion about content moderation, especially the question of what is next for social media platforms. Affected companies, both the social media platforms themselves and the partnering BPOs that supply discounted labour from around the world, are involved in seeking the adequate response to this question and civil society's call for change. This chapter delves into how, instead of diverting away from the brutal racialization of the vulnerable, the Western tech companies and their global partners have worked to retrench the racist map of the colonial condition by furthering and deepening their labour-seeking behaviours in the global south. Specifically, the chapter investigates Facebook's 2019 effort to trial AI

systems in their content moderation process, which came about as a part of their COVID-19 pandemic response, and looks to this case study as an illustrative instance of the shortcomings of AI solutionism, both in its rhetorical form and practical application. I hope the case study clearly points to how human moderators are envisioned and conceptualized both as a stop-gap measure that is temporary and as a more permanent fixture of the content moderation process, as a part of the AI systems.

The growing problem called "social media"

Since the early days of its coverage, journalists and civil advocates have engaged in the work of exposing the unspoken, hidden trade of content moderation. They have publicized direct testimonies of the contract workers from around the world and have worked through the irony of serving as the limited avenues of communication for former and current content moderators, many of whom hid behind anonymity for fear of reprisal from their former employers and the potential legal repercussions of breaking their NDAs. These testimonies have detailed the unrelenting working conditions and brutal performance metrics that content moderators faced, and provided a graphic account of their exposure to traumatic content (Gilbert, 2019; Gray, 2019; Hern, 2019; Newton, 2019; O'Connell, 2019; Roth, 2021).

Insider testimonials from within the ranks of Facebook's software engineers, product managers, and policy experts (representing a class of workers that Christian Fuchs refer to as "direct knowledge workers") join this public scrutiny in expressing their individual and collective dismay, some leaving their jobs at Facebook because they no longer felt that they could make meaningful change within the company (Canales, 2021a, 2021b). Some of these internal pressures have burst out into public attention in the form of whistleblowing, resulting in expansive exposure of internal documents through media outlets ("The Facebook Files," 2021).

These documents have publicized the various failings of Facebook as well as the platform's prior knowledge of the failures, and have brought greater public attention to the ways in which social media platforms mediate social harm. These reports and other public testimonies have also trained the attention of nation-states that have convened legislative hearings on the broader societal impacts of these platforms (C-SPAN, 2021; Foxglove, 2021).

At times, the response from Facebook and its BPO partners to these public pressures have been lackluster, falling short of providing full and adequate mental health support to its workers or otherwise simply asking their new hires to sign an acknowledgement that they may develop PTSD and other mental health disorders by working the job (Perrigo, 2022a), thus merely addressing the symptom and not its root cause. Facebook has also responded by publishing more of its internal data in what they call "Transparency Reports" that detail "how [they] enforce [their] policies, respond to data requests and protect intellectual property, while monitoring dynamics that limit access to Meta technologies." The "Transparency Reports" have included "Community Standards Enforcement Report" which specifically tracks how each content category performs in their moderation and removal from the platform. However, these reports have come short of explicitly acknowledging the harrowing conditions and nature of labour that goes into moderating the contents, or providing solutions to their globalized distribution, and overall have left the impression that their data releases have been a form of public relations exercise.

The solution of more: bodies and machines

Amidst this public scrutiny, the specific response that Facebook and its parent company Meta have provided apropos content moderation and its labour arrangement has two parts: (1) as the platform becomes bigger, attracting more users and more harmful content, Facebook will

respond by adding more content moderators to handle the increased traffic and its expansion into various parts of the world; and (2) because this process of scaling up cannot be done infinitely, Facebook will increase automation on its platform with the aim of lessening the workload put on its human agents. Collectively, these two prongs seem to paint a picture of a company looking to lessen human involvement in its process. However, I argue that they replicate, instead of denouncing, and even intensify, the current content moderation process's reliance on cheap, available, and suitable labour from the global south, and that its infrastructure, by virtue of being automated, continues to involve racialized bodies of its globally outsourced labour force as the platform's sensing and coding mechanisms of trauma.

Solution 1: Scale-up and hire more

On May 3, 2017, Mark Zuckerberg posted on his Facebook page the intention to hire 3,000 more content moderators to tackle the "spate of violent videos that [had] been shared to Facebook in the past month," in part prompted by a recorded murder that was posted on Facebook (Wagner, 2017) and allowed to stay on the platform for two hours before it was taken down (Steer, 2017). This move almost doubled the number of content moderators that Facebook was employing at that time, which was about 4,500 (Zuckerberg, 2017). On October 31, 2017, at a US Senate Judiciary Subcommittee hearing on Russian interference in the 2016 election, Facebook's general counsel at the time, Colin Stretch, responded to the Senate's inquiry on how Facebook expects to safeguard the internet safety by noting that they have about 10,000 people "who are working on safety and security generally and [that Facebook is] committing to investing more and doubling that number by the end of 2018" (Block & Riesewieck, 2018). In 2018, Facebook's then product manager for Civic Engagement, Samidh Chakrabarti, commented on this as well, writing for the company's blog and specifically mentioning the hiring of 10,000

more people as the solution to the problem of foreign interference and the spread of misinformation.

Even when the COVID-19 pandemic made operations difficult, Facebook continued to focus on managing the scale of its workforce and responded by moving its full-time and contractual employees off-site to be able to work remotely. On a call with the press, Mark Zuckerberg noted that this move, in part, involved moving "most types of content that needs extra attention" from the contractual employees to the full-time employees and of increasing the labour pool to be able to address this shift (Zuckerberg, 2022). Content relating to suicide, self-injury, child exploitation, and counter-terrorism were involved in this move, as Facebook felt it was difficult to provide meaningful and continued mental health support to a workforce that was largely remotely distributed. Beyond making a tacit acknowledgement that the work of content moderation requires a more robust support structure than the ones available to the subcontracted employees in the global south, these tactics of crisis management mirrored the tremendous growth of the company that occurred as a direct response to "Facebook's business strategy of relentlessly pursuing user growth in an effort to please investors and the advertisers that are the company's paying customers" (Barrett, 2020, p. 4).

At the same time, Facebook and its BPO partners did not fully engage with the underlying logic of the network that disproportionately locates these labours in historically marginalized and violently impacted parts of the world, nor did they address the way in which the work itself continues to harm those employed and suppress those who seek out public support with uses of business-legal instruments like the NDA (Arsht & Etcovitch, 2018). The attempts at scaling up content moderation has not only come with naturalizing the platform design that fundamentally leads to the traumatization of its workforce, one that responds to the

call of the imperial logic to make certain bodies more available than others for the damaging labour, but it has actively sought to continue and amplify the magnitude of harm that tech companies process through racialized bodies for its own exponential growth and profit. To grow the platform has meant growing the body of the workforce that can process more and more contents and filtering out increasing levels of harm, and with Facebook's growth looking beyond the confines of its established markets, its stated promise to scale its content moderation process to respond to the increasing diversity of language and cultural context has come with the implicit indication that they will spread the harmful work around equally diverse parts of the world.

Solution 2: Bring in the AIs

However, these measures of scaling up content moderation by simply hiring more people were proffered with a hint of caution that they would not be lasting solution to the rising tide of problems that social media platforms face. When Facebook hired 3,000 more content moderators to combat the rise of violent videos and livestreams on its platform, for example, Zuckerberg noted that their solution to the problem may be imperfect: "No matter how many people we have on the team, we'll never be able to look at everything" (Wagner, 2017). Its efforts to scale up simply cannot be a linear, one-to-one solution, especially with network traffic growing at an exponential rate across the globe in various languages, socio-cultural contexts, and in differing degrees of market penetration.

In context, then, proper scaling of content moderation through greater employment has always been considered a stop-gap measure, a short-term solution to the more intractable problems of the growth and spread of the social media network. Companies like Facebook have thus envisioned and worked towards eliminating the human factor from its content moderation process. In the human's stead, the longer-term project imagines machine-driven algorithms and

artificial intelligence technology would take over the role of identifying, sorting, and acting on the contents that are deemed impermissible from its platform. No more human content moderators (at least not at the current scale of operation) but an even faster and more efficient pace of moderation in future. Or so goes the hopeful rhetoric. A variation on this future imaginary has also included a human workforce with a reduced footprint that is augmented and supported by AI systems, thus making their work easier, faster, and more expansive than before, despite the actual reduction of human agents in the process (Meta, 2022b).

In part, this vision was trialed at the outset of Facebook's pandemic response in 2020. With a caution that the machine intelligence may make more mistakes than their human counterparts and otherwise struggle to properly discriminate between "legitimate and illegitimate content" (Bond, 2020; Facebook, 2020; Scott & Kayali, 2020), Facebook deployed an automated content review process, which came in concert with other COVID responses, including sending its content moderation team home to work remotely. Facebook's notice to the public on March 16, 2020, however, included only a very simple and quiet note that they have been "rapidly growing content review teams and expanding [their] machine learning capabilities" (Facebook, 2020), as well as omitting how only 72-hours prior to the blog post announcing its "decision" to send contract employees home, "the entire metro Manila area of the Philippines had just gone under quarantine" (Roberts, 2020).

Roberts has described the tool of the machine-driven content review process as a "blunt" weapon, a tool that can be scaled up in ways that human moderation cannot (Roberts, 2020), but such scale comes at the expense of losing the nuance and precision offered by its human counterpart. The aftermath of Facebook's machine-learning driven content review process regime in 2020 seems to have affirmed this critical trade-off. In October of same year, Scott and

Kayali (2020) reported that in the second quarter of 2020, Facebook "roughly doubled the amount of potentially harmful material they removed in the second quarter […] compared with the three months through March," but that this increase came at the cost of "dangerous and possibly illegal material [being] more likely to slip past the machines." In Syria, several activists accounts that recorded and archived evidence of war crimes were banned from the platform. The banned account holders had little recourse for appeal, leaving their often time-sensitive efforts in jeopardy. News articles and health information concerning the new coronavirus were also taken off the platform, flagged by the AI system as inappropriate. Meanwhile, child exploitation and self-harm were two of the categories of content that the removal process failed to catch, and subsequently led to a steep drop of "at least 40 percent in the second quarter of 2020" in its deletions from the platform due to "a lack of humans to make the tough calls about what broke the platform's rules" (Scott & Kayali, 2020).

Facebook's response to such failures were to bring humans back into the process. When Facebook's "beta test" with AI systems failed to yield the result that they wanted, content moderators were recalled back into the office, so that more of them could be brought "online." The return to the office was mandated despite the continued dangers of the pandemic, and the workforce of the global south had to bear the brunt of this danger as they were recalled amidst multiple COVID outbreaks (*Open Letter from Content Moderators*, 2020). Subcontracting entities like TaskUs, at the behest of their client companies, led the charge of bringing back the workforce to fill in the need of moderating harmful content accurately and efficiently, at times operating within a complex sequence of receiving input from or providing data to the AI systems that were meant to aid them (Chugg, 2021). On a path to broader AI adoption, human involvement seems crucial in a way that is more than perfunctory or temporary. As their required

presence in the high-risk content moderation categories within a high-risk work environment have demonstrated, a human being's capacity to sense specific harm continues to be a critical function within the broader socio-technical mechanism of content moderation, and the fulfillment of this function will be sought no matter the cost.

The Real AIs in the solution

This kind of recall may present the problem of AI as being a technology that is incomplete or inadequate *at the moment*. However, this is a formulation of the problem that assumes and indeed relies on an imaginary of distant future when such issues have all been sorted out through the inevitable march of progress. What stands at the root of Facebook's recall is the material reality that these two – the human reviewers and the machine intelligence – are intractably interlinked and are constitutive of the overall infrastructure of the content moderation that many companies employ today or are otherwise seeking to transition towards in the future. Despite this complex picture, there continues to be a tendency to imagine AI systems, especially in the context of content moderation, as an independent process separate from human input, a sentiment perhaps most succinctly echoed in an interview with a content reviewer who spoke to Politico: "[The COVID-19 pandemic] just fast-tracked something that many of us saw coming…. We are only here until the AI tools can catch up" (Scott & Kayali, 2020).

Scholars have noted the rampant solutionism in Silicon Valley, where complex social situations are recast "either as neat problems with definite, computable solutions or as transparent and self-evident processes that can be easily optimized" with only the right algorithms in place (Benjamin, 2019; Morozov, 2013, p. 5; Taffel, 2018). The world of content moderation and the social media companies themselves also participate in this broader trend of techno-solutionism, leaning heavily on machine learning algorithms' potential and their

ostensibly inevitable capacity to process huge amounts of data to output knowledge about the contents being transmitted through its platform (and in part about the platform itself) at an incredible speed. They at times attempt to forecast the future where content moderation's human cost problem is made a thing of the past, by pointing to select successes today, like that of the "90 percent" detection rate that Facebook touts on its webpage that describes how their technology detects violations (Meta, 2022a). However, this type of solutionist outlook and reduction of design principles to simple percentages and prevalence rates may continue to exacerbate the extant problem of human content moderation by promising unrealistic expectations of what AI can do and by presenting a false or incomplete picture of what AI technology currently is. Evgeny Morozov, when describing the AI solutionists trends of Silicon Valley tech companies, cites design theorist Michael Dobbins to note that solutionism presumes rather than investigates; it reaches for answers before the questions have been fully asked (2013, p. 6). The AI solutionism of Silicon Valley, and of Facebook, seems to presume that the evolution of machine learning is inevitable, as is its succession over the current manual process. Even when considering the two parts of the system not to be mutually exclusive, the "AI solution" seems to imagine that a harmonious relationship grounds the relational dynamic between human agents and machine learning systems, instead of acknowledging the complexity of their relationship, which is rather racially inflected and subordinate.

Raced ghost in the machine

To fully understand the human-machine relationship that AI technology envisions, we must understand that human content moderators and other digital data workers of the global south not only actively participate in the knowledge production process that AI systems are purported to drive, but that they in fact form the foundational programming protocols of the AI

systems in the first place. Beyond working as a temporary stand-in for the inevitable takeover by the mature machines systems, human agents provide significant "computational power" and serve as machine-learning systems' interface to the material reality by processing "data" through their own bodies, thus preparing datasets that serve to educate and "train" the machine systems.

Creating datasets for AI system, in its most basic form, requires human agents to review images, for example, and classify digital objects by labelling them in a machine-readable format. The drudgeries of this monotonous and repetitive work is referred to by Roberts as representing the "work of the new digital assembly line: rote, repetitive, quota-driven, queue-based, automated tasks relying on mechanization and rationalized management practices and serving up metrics to assess employee productivity and efficiency" (Roberts, 2019, p. 92). Other scholars have also noted how a disproportionate amount of AI knowledge sets are generated by racialized labour forces that go unnoticed (Benjamin, 2019; O'Neil, 2016). Marte-Wood and Santos point out that this imagining of content moderators' labour (as a form of "digital assembly line") helps rationalize the sheer volume of the quota that they must meet as a part of their employment (Marte-Wood & Santos, 2021, p. 107). For me, this way of understanding their labour helps highlight a significant fact of AI systems in modern-day communication infrastructure: the knowledge that gets generated by AI systems is not entirely artificial or "inhuman," nor is it the product of "direct knowledge workers" that designed and "programmed" the AI systems and protocols. Knowledge generated by AI systems is derived from much more bodily input than we conventionally imagine.

In the case of Facebook, human workers are fundamental in generating various datasets that their AI systems require for training, for generating knowledge about the digital and digitized artefacts that go through its platforms, and for providing corresponding outputs (i.e.,

deleting or permitting any of these artefacts). This training process takes place by looking at the actions that content moderators have taken, translating them into programmable data that can provide AI systems with a roadmap for how to act when dealing with similar contents in the future; or it can take place by getting human agents to prepare a dataset and otherwise manually labelling digital artefacts according to corresponding community guideline and platform polices that it purportedly violates (Meta, 2022a). These data relations created by human agents are then inputted into the AI system through a “human-technology feedback loop” that animates the path towards greater automation (Meta, 2022a). In this process, the human agents play the instrumental role of making Facebook’s algorithms be what they are, and in keeping the systems up-to-date on the types of harmful contents that permeate the network, they help it stay relevant and current. In other words, the efficacy of Facebook’s AI systems comes largely from the human inputs of the “indirect knowledge workers” of the global south that they employ to process, label, and codify in machine-readable format (Meta, 2022a).

In the context of BPO operations, these workers that comprise the AI sit alongside the workforce that moderate content, sometimes quite literally, in global south locations where that work has been outsourced. TaskUs, for example, structures its services in three distinct channels: (1) Digital Customer Experience, (2) Content Security, (3) Artificial Intelligence Operations. TaskUs describes its Artificial Intelligence Operations as “principally [consisting] of data labelling, annotation and transcriptions services performed for the purpose of training and tuning AI algorithms through the process of machine learning” (Chugg, 2021). TaskUs also explicitly notes that this work of providing data annotation is not only for providing a starting point for AI training, but is also necessary for the maintenance of its AI system which “must be refined through ongoing manual training” (Chugg, 2021). In effect, one of its advertisements notes that

their AI services enable companies to "scale training data programs efficiently, effectively – and excellently" (TaskUs, n.d.), further illustrating that the often-noted traits of artificial-training technologies are, in this case, coming from the human workers themselves.

This organizational division and simultaneous closeness between content moderation and AI data labelling continue the strategy of divide-and-conquer that has come to form the basis of ongoing corporate efforts to bring labour costs down. In her interview with commercial content moderators, Roberts noted how the content moderators were being compensated less than their colleagues that worked on a "comparable voice account" (Roberts, 2019, p. 188). Despite their "specialization" and focus, and even though their respective labours involved managing the type of emotional response of the American users and callers, the content moderators' labour was not valued the same as their voice-account colleagues. This value difference, Roberts noted, varied even within the same company, depending on their "portfolio" and "account," and more so between different BPO companies that the moderators worked for. This type of value difference also fuels Western companies' efforts to "automate," with the further division of the care work that goes into mediating content now providing an additional rationale for discounting the labour value of the workers being hailed into the overall mechanism of social media infrastructure. In one instance, an individual working as a micro-tasker for a subcontracting firm was noted as being paid 30 cents per hour for labelling datasets (Rigaud, 2019). Reminiscent of the way Chinese railroad workers were brought in as a cost-saving measure, we see the way in which racialized content moderators are forced to engage in a dynamic web of price differentials that ultimately benefits the tech companies and their BPO partners.

The embedded nature of human input in the machine learning process has a particular implication for the moderation of harmful content. Monitoring harmful content employs human

agents as its technological filter, leveraging the expected affective responses that the human agents generate in response to being exposed to contents that are suspected or flagged as harmful. Trauma, in this case, is not an unwanted by-product of the whole process but is the very thing that the whole process needs to be able to process and operate accurately. Content moderators (and, in the case of AI services, the human annotators) and their capacity to be traumatized is precisely the thing that makes their labour valuable to the overall industry. Machines are unable to feel trauma, but humans can. In the early days of content moderation, when the work was done in-house by a handful of young American cohorts, who landed their jobs first thing out of university, a similar requirement of "bodily response" to the contents existed: "'We were supposed to delete things like Hitler and naked people,' recalls Dave Willner, a member of Facebook's pioneering content moderation group. He and his colleagues were told to remove material 'that made you feel bad in your stomach'" (Barrett, 2020, p. 7). This requirement of bodily response to the contents that they moderate and annotate continues in the contemporary mode of outsourcing the work, which has now become the infrastructural basis of the artificial intelligence solution of social media platforms as. In other words, the human agents that annotate for AI services are in fact "encoding" into the dataset trauma and "signs of bodily stress" (Jereza, 2021, p. 930) that come through their experience of a wide range of affective states, often harming themselves in the process, translating their own trauma responses in machine-readable form.

The case study of Facebook's 2019 trial of AI systems in its content moderation process, as well as the recall of its human workforce and the way it converted some of its more sensitive and harmful content portfolios to fulltime employment, shows us how they viewed their workforce as a critical part of AI. It demonstrates the constant requirement of human agents in

the overall process of AI systems, not only as its starting point, but also as its trainer, failsafe, and maintenance support, as well as a media filter that senses harm, experiences it, and codifies the experienced trauma as a machine-readable knowledge. In this way, the human digital workers, including the content moderators, are not only producing "immaterial knowledge" about the platform and its content through their analytical work, but they fulfill their role through their "care" and through their bodily presence.

Code of care

As noted in the previous chapter, the suitability of the Filipino workers in this role and in their position within the grander infrastructure of social media platforms are enabled by the postcolonial legacy of US occupation in the Philippines. Marte-Wood and Santos note that United States has a long history of "developing Filipino care workers, as seen in the OFW[6] laboring as domestic workers, nurses, caregivers, and teachers" (Marte-Wood & Santos, 2021). Concentration of content moderation work, and indeed the global outsourcing of the menial and underpaid digital labour, to the Philippines continues this history in a new form, extracting care work from the region and its people as they "work in an industry that extracts their vitality as biocapital to safeguard digital lives in United States" (Marte-Wood & Santos, 2021, p. 106). Marte-Wood and Santos see this process as being a form of "remediation" of the extraction of care-labour, now made possible by the "disassembly and distribution of Filipino labor via the BPO industry, after which Filipino care is made akin to code" (2021, p. 106). This kind of remediation of labour extraction then can perhaps be seen as parallel in the relationship between the history of Chinese railroad workers and the Filipino content moderators, but also in the way in which the historical formation of call-centre outsourcing has evolved to include the "omni-

[6] Also known as "Overseas Filipino Workers."

channel contact centre support" structure that now expands the types of labour done overseas to include non-voice accounts like content moderation and data labelling (Simplr, 2019).

There are many different global sites that seek to participate in this economic exchange undergirding social media platforms' services and profitability. However, the predominance of Filipino labourers on the global market carefully echoes the historical legacy that has fostered, by force, "a proximity to American culture" and the overall "Filipino/American relatability" through "benevolent" educational system that has become a foundational aspect of the "value-added aspect of the Filipino BPO" for US-based tech companies (Marte-Wood & Santos, 2021, p. 105; Roberts, 2019). The affective proximity created by such history has meant that that "wrong feeling in the gut," which formed the basis of rudimentary content moderation policies at the inception of Facebook in 2008 and were exercised by the young American college graduates as their rule of thumb for knowing which content to delete is now utilized wholesale by Filipino workers as a systemic and fundamental logic of how content moderation is supposed to work, at scale but with lower pay (MicroSourcing, n.d.-b).

Similar to understanding "resilience" as more than an individual quality and as a process of categorization, "intimacy" in the context of BPO services must be understood in relation to how racialized bodies are prepared for the gruelling digital labour that helps direct flows of capital and communication globally, and as the rationalizing logic that makes them a natural machine part of the platform's AI future. By finding an "intimate" labour subject in the figure of Filipino content moderator, Western tech companies and their BPO partners are in fact constructing a kind of avatar of an American labour force, one that can better take on the task of sensing trauma by being traumatized. In the figure of Filipino moderators and data labellers, American tech companies like Facebook are finding a figure and body that can pre-emptively

sense content that may harm its userbase or its platform value. In this way, they have come to comprise the basis of the platform's machine-learning systems, content moderation, as well as its user experience design.

This kind of availability has also buttressed the profitability of select Western tech companies and multinational BPO outfits: "The rise of Filipino BPO work generally, and commercial content moderation specifically, signals an important shift in the decades-long, state-sponsored 'expropriation of care'" (Marte-Wood & Santos, 2021, p. 111). For Marte-Wood and Santos, commercialized content moderation stands out among other "older remediated forms of Filipino care," as while the older forms stand relatively "untransformed by technology," CCM represents a newer iteration that does "the most thorough technological recomposition of Filipino care to date" (2021, p. 110). Through these new digital intimacies, racialized and gendered bodies undergo "virtual migration" that extends beyond the historical migratory patterns of a physical move overseas, thus remediating "form of Filipino care historically specific to the rise of American platform capitalism" (Marte-Wood & Santos, 2021, p. 110) from "older networks of global care chains, state-sponsored labor export, and economic remittances."

The paradoxical mobility of the racialized bodies is in part constructed and capitalized by the Filipino state, which has acted as labour broker to play "a critical role of producing, distributing and regulating Filipina as care workers across the globe," increasingly privileging Filipino care's digital distribution over its physical export (Marte-Wood & Santos, 2021, p. 111). In March of 1994, the Philippines connected to the World Wide Web for the first time, and while official accounts depicted the watershed moment as "good-faith collaboration between industry, state, and academe, the connection concretely marked the moment of the Philippines' internet traffic being routed through the United States, a moment that Marte-Wood and Santos argue

recalls “the uneven, century-long distribution of Filipino care work and intimate labor to the United States” (2021, p. 111). It was the internet protocol that established connection via SprintLink and routed the Philippines connectivity through California, but it was also the signing of Republic Act 7925 (also known as the Public Telecommunications Policy Act of the Philippines and RA 7925) into law that ensured the use of this global connectivity for the purpose of making Philippines labour available by creating a route of “easier foreign direct investment” (Marte-Wood & Santos, 2021, p. 112). RA 7925 sought to achieve the contradictory aims of meeting the needs of the Philippines people by keeping them in the country while taking advantage of the export opportunities under a fast changing telecommunication, stating as much as that “the private sector shall be the engine of rapid and efficient growth in the telecommunication industry” (*R.A. 7925*, 1995). Marte-Wood and Santos note that “More than any other legislative policy, RA 7925 helped transform the Philippines into a global hub for outsourced and offshored information and communications technology (ICT) services” (Marte-Wood & Santos, 2021, p. 112). In this way, Filipino bodies have in turn become a part of the internet protocol and digital infrastructure that provides a global userbase (especially that of American and other English-speaking users) access into the moderated media environment of social media platforms.

This national strategy of hanging on to its labour force domestically has complemented the creation of economic zones that provide tax relief to foreign investors and attracting American tech companies’ labour needs by providing labour resources in disassembled form, a process that dismantled bodily attributes into “eyes scanning an image in few seconds, into fingers clicking mouse buttons twenty-five thousand times a day,” and as a labour force that works efficiently, “supposedly insusceptible to the traumas experienced by their white U.S.

counterparts" (Marte-Wood & Santos, 2021, p. 113). Tarleton Gillespie notes that content moderation itself is, in many ways, "the commodity" that the platforms offer (2018, p. 11). To be able to offer this "commodity," then, the Filipino workers' bodies are made mobile within the transnational circuit of care labour and are racialized as a body that can be violently opened up, carved out, and ultimately made a suitable body for taxing and divisible digital labour.

Conclusion

The dingy darkness of the office space that opened the first chapter of this thesis and the documentary *The Cleaners* (Block & Riesewieck, 2018) may no longer accurately represent the contours of the labour and its working conditions that many content moderators find themselves in today. TaskUs, for example, touts that "imaginative, inspiring, and 'Instagram-able'" (Chugg, 2021) office spaces, mirroring many of the Silicon Valley tech start-ups, are available for its employees. Its employees, TaskUs says, may access its many on-site amenities (gym and child care on its "creatively designed sites") and benefit from their wellness monitoring (providing "healthy meals" at its on-site cafes). TaskUs claims that their expansion plans set it up to be an ideal workplace for many Philippines locals (Chugg, 2021).

However, these bright, modern workspaces and their corporate "benevolence," as well as the diversifying labour portfolios of the BPOs that bring in more automation, do not undercut the rhetorical power of the imagery of dark office spaces. What these dark spaces provide is an illustrative throughline between the imagery of darkness and the violent reality of BPO labour by pointing squarely to the disembodied labour that the global south workers continue to do, as corporate valuation increasingly becomes tied to the way in which Western tech companies and their global partners can forcibly cleave apart the "historical and material specificities" from Filipino workers' bodies in an effort to network its media infrastructure. The technique of

dramatization that *The Cleaners* uses also provides what Lisa Parks refers to as the "infrastructural perspective" of what it is like to be enmeshed in the social media infrastructure as its filtering resource. It does so by reflecting the embodied experience of the workers whose work simultaneously involves connecting users while being themselves socially isolated in their traumatizing routine (Parks, 2015, p. 363). Otherwise put, the dramatization helps the viewers to understand that the proper functioning of social media platforms is contingent on labour conditions that create and exacerbate social isolation amongst their workers, work-routine monotony, and constant exposure to traumatizing contents and/or audio-visual stimuli.

In this way, the digital transformation of the racialized labour has meant that the new forms of mobility of the body have integrated racialized bodies into the broader media environment that continues to prioritize generating value and capital for the West. The specificity of racialization has worked to make Filipino bodies "less autonomous and more automated" and prioritized the workers' capacity to care (i.e., to empathize and to be traumatized) as a crucial part of their role as an "information machine" that identifies harm in content, applies (platforms' community code of conduct), enforces (by deleting or permitting contents under review), and ultimately monitors the "health of the platform" (Marte-Wood & Santos, 2021, p. 106). There has been an ultimate reduction of the "complexity of Filipino empathy to a finite set of possible machine expressions, effectively binarizing care into code" (Marte-Wood & Santos, 2021, p. 115), and this type of machining the race echoes myriad observations made by scholars that point to how this mode of networking to the greater web (or "the net") has always presupposed dispossession of the racialized of their own bodies and sense of self:

> It is not simply, then, that some have had access to disembodiment and others have not but rather that some have never had a body—in the sense of an integrated whole whose skin is seamless and unmarked—from which to abstract in the first place. The fiction of the disembodied citizen depends on this other disembodiment (W. H.-K. Chun, 2006, pp. 135–136).

TaskUs claims that their provincial expansion strategy to locate its offshore sites away from the densely populated major "tier 1" cities and to rural areas "near where employees live" will help contribute to a "superior employee experience" by helping to reduce commute times (Chugg, 2021). But in truth, their business strategy seeks to benefit from reduced competition for talent (i.e., by being the only viable employment option) and from paying lower salaries and from exploiting the lower rents and cost of living standards in those regions. This kind of "benevolence" of the tech companies and their partners hides the underlying corporate greed that pits different groups against each other to bring costs down, whether it be between companies, countries, regions, or races. Looking at this exploitative picture makes it hard to unquestioningly applaud corporate celebration of "higher employee retention," whether it is achieved by making people's commute more palatable or by providing resources to garner more resiliency in their workforce or by reducing the workload through automation, for what undergirds their corporate planning is the profit they stand to gain by reducing new populations in the Philippines and global south to beings "for the captor" (Spillers, 1987, p. 67).

Seeing care as "akin to code" denotes the specificity of the 21st century global "knowledge economy" that builds on transformations enabled by interconnected networks, globalized labour outsourcing processes, and the "logic of disassembly." There is a paradox of mobility that stays the racialized bodies while simultaneously making them the conduit for the flow of digital signals, moving them when necessary, but also making their values in the flux of a global market, and considering their bodies and labour programmable in the unequal labour distribution, which overall reproduces the "new assembled, modular, and distributed supply chains of global capitalism" with the material resource of the racialized bodies (Marte-Wood & Santos, 2021, p. 113). This construction of the global media channels has meant "terraforming"

the global south regions in the shape of the West's capital needs, thus helping to quell the media environments through control and violence exerted against a historically dispossessed population. This history continues to echo the previous mode of colonial violence that also moved racialized bodies, displacing them and relocating them to American shores, so that they could become the tool for carving out mountain paths of its railway networks. These bodies so moved not only helped to construct the instrumental infrastructure through which the American sense of self and its economy flowed, but they also moved imaginaries of the future, which in the late 1800s was one of American unity and imperial emergence. Today, the future echoes the past, with companies like Facebook heavily leaning on their own potential to connect the world, establishing their foothold in the globalized digital economy in the process.

In this sense, it is clear that much of these companies' publicly shared visions of what a future for their platform might look like, and how much of that future is dependent on the notion of "fully fledged AI technology," is not an idle imaginary of a hopeful future. Companies like Facebook silently presume continued human input on an ever increasing scale to achieve their hypothetical and amorphous future, and it must be noted that their imagination of an automated platform future is one that continue to hail racialized labour into its infrastructure, this time at greater discounts and with fewer labour protections. As social media companies like Facebook gesture to a machine-driven future, the public must be wary and recognize that such promises hide the dreary makeup of its platform's present material infrastructure.

CHAPTER 4. MEDIA OF THE FUTURE

I'm different from what I'm before.

-- unnamed content moderator, *The Cleaners*

The next thing

This chapter brings my research to a close by looking to the future of content moderation, both as a deepening practice of globalized labour outsourcing and as a foundational infrastructure of the new digital economy of "Meta"-verse, of which Facebook and its partners seem intent on playing a rather central role in the years to come. This kind of emergence and dominance does not come without its critics, and I refer to several of them to underline public's growing awareness and wariness of the dominant social media platforms' impact on today's world. Through this research I have sought to develop and add my observation that a fuller understanding of social media networks must include the recognition that their infrastructures are materially made up of the bodies and inputs of racialized workers from the global south, whose labour and humanity are violently cleaved from them as they mediate our daily connections, filling the pockets of select corporate actors.

Part of this understanding has come by looking backwards at another history of construction of a communication network that made violent use of racialized labour. The affinity that exists between the racialized labour of the Chinese railroad workers and the Filipino content moderators helped me demonstrate the historical continuities that contributed to forming the Colonial mediality of our contemporary media environments. To socially connect is not such a socially neutral project, and by challenging the neutrality of the network through which we communicate daily, I call for an ethical encounter with others that challenges the naturalized notion of race and look to Beth Coleman's call for careful, ethical encounters with "race as

technology." As Coleman notes, "if race possesses no value without context, then we must choose to act courageously when faced with oppression – our own or somebody else's" (Coleman, 2009, p. 181). Lastly, I close the chapter by outlining how such ethical encounters with race as technology have occurred in the history of communication networks by looking at the history of Chinese railroad workers' intimacy with the Umatilla people, and how they are currently happening by looking at the ways in which content moderators themselves are resisting their own racialization and mediatization through labour organizing, whistleblowing, and waging legal battles with their employers.

Future looking

Most recent estimates and projections about Facebook, as a social media company and one of the largest contractors of content moderators, present a staggering picture of a company whose exponential growth has persisted since its inception, despite its recent political turmoil in 2016. Between 2016 and 2021, Facebook has invested more than 13 billion dollars in efforts to "ramp up its security measures," and has more than quadrupled the number of employees working on content moderation from 10,000 to 40,000 (Beals, 2021; Meta, 2021c). This ramp-up reflects the steady growth in Facebook's userbase and revenue since 2016 (Statista Research Department, 2022a, 2022b). The accelerated growth of social media companies like Facebook (as well as other tech-enabled market segments like fintech, digital health, and streaming services, to name a few) is what BPO companies like TaskUs index their own projected growth against. BPO companies and their current infrastructure (which expands daily) continues to help externalize unwanted cost (of attention, risk, and capital) of scaling and even starting up Western tech companies:

> As technology companies scale, they must dedicate resources across product development and operations. However, they often *lack the physical capacity or desire*

> *to develop operational infrastructure* internally as they focus on growing their core offerings. As a result, we believe technology companies are increasingly willing to outsource at earlier stages of their lifecycles and are driving outsized growth within the overall outsourcing industry (italics mine) (Chugg, 2021).

With projected market growth in Digital Customer Experience, Content Security, AI services at Compound Annual Growth Rate (CAGR) of over 20%, 40%, and 20% between 2018–2021, 2016–2021, and 2020–2024, respectively, the future of content moderation and the global outsourcing of the harsh labours involved (both in actually moderating content and in helping to form parts of the AI services that are meant to replace them) seem reliant on equally growing the human labour pool in regions that can supply the labour and meet the demands of the world market at reduced cost. TaskUs's above-mentioned regional expansion plan also suggests that this kind of labour need will be met by making more of the global south (specifically the Philippines) available for the digital economy going forward.

Over the course of this research, reading various growth statements about the future of social media, both about the industry's historical growth and future trajectory, have often made me feel as if I was bearing witness to a tremendous tide. The incredible scaling of the extant mode of social media and its evolutionary expansion into the "metaverse" seem to signal the tech companies' unshakeable confidence in their own future. They seem so certain that they will not only grow and grow, but that their future is inextricably linked to the nebulous notion of "public good," that their success means enriching the world through global connections. Facebook's ongoing capacity to consolidate various social media outlets (like WhatsApp and Instagram, for example) under a single business entity, Meta, have also signalled for me the tremendous power that has been allowed to centralize in the hands of a select few since the early days of Web 2.0. In these ways, their recent rebranding's focus on the "metaverse" and on bringing as much of our

sociality, its play and work and living, all under its services seems rather to signal the emergence of a “Meta”-verse.

Of course, the public no longer seems to accept these bold claims of universal connection (and that such connections lead to social good) at face-value. There is growing skepticism around the promises of tech companies, and researchers and journalists have continued to explore and expose the inner workings of social media platforms and their networked relations. The investigative works of those that brought to general attention the hidden workings of content moderators in the first place (A. Chen, 2014; Newton, 2019) are being continued by those that explore how the unequal web of globalized labour distribution expands into more regions of the world and in matching the increased expectations of the tech companies, and of the userbase, for the permanently online services (Perrigo, 2022a; “The Facebook Files,” 2021). There have been leaks, whistleblowers, and legislative meetings around the world. Through these efforts (which have had various degrees of success and efficacy), the public have come to know more about the platforms that we rely on for today’s communication, but at the same time these disruptions have also demonstrated that tech companies are more than capable of and indeed do exercise as a part of their operation obfuscation of the public’s attempt at knowing more about one of the biggest knowledge machine of modern times.

What I seek to add to the chorus of critics and observers of social media’s growth is a reminder that our ongoing scrutiny of the technology and medium of social media must include the understanding of race as a fundamental *technological (as well as technologized) and material component* of social media. Through this research and especially through this chapter, I am reiterating the importance of seeing beyond the contents (and the harm of the contents) of social media networks as definitively marking its flow. Instead, I shift the focus to the racial and

technological makeup of the infrastructure, and how such materiality impacts the mediation of contents, harmful or otherwise, in its requirement of violation against racialized bodies. Giving attention to the materiality of the infrastructure also means that in doing we can see the incredible valuation of social media companies and their technologies for what it is; not only as a measurement of the platforms' capacity to attract large numbers of users and transmit incredible volumes of content, but also as being driven by the instrumentalization of certain classes of bodies as a part of their material infrastructure of digital delivery. All these considerations, I believe, ask us to question the received wisdom about our communication networks as being neutral mediums. And when disbelieving their neutrality, we may even begin to ask "how else can we connect?" with hopes of understanding how we can ethically resist such violent technologization, instead of acting as subjects wielding the technology (of race) as our own extensions, a prosthesis for ourselves.

Race as technology

Beth Coleman (2009) understands the importance of asking these kinds of questions when she considers "race as technology." Her invocation moved race away from the "biological and genetic systems that have historically dominated" the definition of race and instead posited that the consideration of race should be a question of "technological agency," and as the way by which "external devices help us navigate the terrain in which we live" (Coleman, 2009, p. 177).

Coleman explains that the technological mechanism of "race" refers not only to race as a mechanism of parts, but rather as something that is constructed to work as a "lever-like" mechanism that functions *systemically* and as "an array of procedures" that creates certain things or people (italics mine) (Coleman, 2009, pp. 190–191). The makeup of the network of social media and the globalized labour outsourcing embraces this quality of technology that brings

together racialized peoples as different “parts” of its technological apparatus in various degrees of subordination to create connections that carry meaning and value.

Here, to solely focus on the way in which racialized individuals have been interpellated as machine-like (e.g., “robotically unfeeling Asians can work on and on and on!”) is perhaps to fall into the same trap of techno-orientalist readings of the 1800s and today. Instead, we must emphasize that the rhetorical and material construction of racialized individuals as a form of technological solution takes place on a *systematic* level, through repeated engagement of social relations and forces, not only as a technological fix or development.

The central thesis of “race as technology” asks us to look at the ways in which networks create and actuate the lever-like mechanism of race to construct, often violently, individuals as classes of people that are read as being disposable and capable of withstanding stupendous amounts of harm and pain. In the context of the digital economy, we see that this class of people is summoned to generate wealth and knowledge for the designers of technology (of race and social media), its primary “contract holders” located predominantly in America. As Coleman notes, “to elucidate the general mechanisms that create or continue oppressive forms of inequality, we must also attend to the networks of power: the material and ideological structure of the machine” (2009, p. 191).

However, Coleman’s contribution of “race as technology,” I think, is not only meaningful for her clear acknowledgement of the severe material output of such machines, whose concrete externalities do indeed erase, violate, and “disassemble” the people so involved: “within the construct, if you die, you really are dead” (Coleman, 2009, p. 183). Rather, Coleman’s bigger intervention is to assert that race, when seen as a technology and a prosthesis, can be wielded for an “ethical encounter with others,” as a “disruptive technology that changes the terms of

engagement with an all-too familiar system of representation and power" (Coleman, 2009, p. 178). In attending to the "ethical aspect of race and human difference," we must understand how race as a "sliding value" is to be wielded, not only by its "designers" but by the very people so technologized, as a means of resistance and as a mode of solidarity.

The resistance

In his writing, Hsinya Huang ruminates on the historical intimacy that budded between the Chinese labourers and various Native American tribes that worked and lived along the expanding railway. Chinese labourers introduced new vegetables, like cabbage, to the culinary habits of the Umatilla people, and "in return" they were fed (Hsinya, 2019, p. 191). Huang reads these mutual exchanges as a material connection between two "classes" of people who occupied the bottom rungs of the American racial system, which in fleeting moments allowed them to exercise their agency to connect with others and their histories in ways previously not possible: "[The Umatilla people] exercise, as Jacques Derrida puts it, their 'sovereignty as host.' Hospitality toward Chinese laborers thus involves indigenous claims to the land as well as the desire to establish a form of self-identity" (Hsinya, 2019, p. 192). Chang notes that the "erasure" of Chinese people or their history, even though systemically attempted, was not so uniform and that "The Chinese railroad worker as a historical figure was never stable" (Chang et al., 2019, p. 23). These stories of mutual connection and recognition, which are orally handed down and are maintained in the tribal archives, fills in gaps left by the dominant narrative, thus countering that history's Colonial mediality and helping to form the basis of counter-histories of solidarity and intimacy for its own sake.

The Chinese workers, at times, also directly resisted the unfavourable conditions of their work. From their refusal to work the Native American lands that were "not yet controlled by

white settlers” (Hsinya, 2019, p. 187) to their notable strike of 1867, which was for better wages and fewer working hours, there are historical records of when the empire’s racial interpellation failed to assert its power. To break the strike of 1867, E. B. Crocker and Mark Hopkins, both part owners of the CPRR, looked to hiring “recently freed slaves as strike-breakers”: “Hopkins reasons, ‘A Negro labor force would tend to keep the Chinese steady, as the Chinese have kept the Irishmen quiet’” (Chang et al., 2019, p. 15). But the freedmen never arrived on the CPRR line. There is no clear record of why they never arrived, but it is nonetheless telling of what those moments of resistance (or failure) meant to the Chinese labourers that sought to resist their own instrumentalization. In those moments, race was not a technology or a tool that unquestioningly abided by the will of the imperial power. Instead, they served as a stark reminder of the agential beings that asserted their ownership of self and demanded better.

These types of resistance can even be found in the contemporary scholarship about the Chinese labourers. Without wanting to present an overly utopian take of the internet and its digital communication, it is worth mentioning that Chang cites the “digital revolution” as one of the major factors that contributed to the speedy and voluminous research undertaken by an international group of scholars. Through digital means, the transnational research effort was able to help fill up the historical gap about the Chinese labourers in a way not previously possible: “The ability to share image, text, comments, and questions electronically greatly facilitated collaboration” (Chang et al., 2019, p. 5).

The construction of the transcontinental railroad was seen by many as a symbolic triumph and an opportunity to heal for a country divided by a “bloody civil war”: “the railroad that bound the East Coast to the West was hailed as an emblem of both unity and progress” (Chang et al., 2019, p. 17). However, my rumination on the history of the Chinese migrant labourers makes it

apparent that such "triumph" was a result of an ongoing negotiation of the racial hierarchy which had been significantly disrupted in the aftermath of the Civil War. Richardson's reading of the post-Civil War South makes it abundantly clear that this was not simply a time of emancipation, but rather a turbulent time of newly emergent racial hierarchy which was just as stringent and vertically divided as the one that preceded it. There was a "new logic" emerging, one that would keep pace with the new realities of modernity, but one that also saw it fit to "blacken" its racialized subject for the profits and benefits of the imperial economy and its need for infrastructure to sustain it (Towns, 2019). The empire that had profited off of racialized bodies, especially Black and Indigenous, was not about to let it go so easily.

However, what the history of transcontinental railway also illustrates is that there is a way in which racialized subjects can "slip out" of their fixed positions within the racial hierarchy, not by some magic, but by recognizing their own "Black mediality" and by reaching out to others who were similarly racialized, an extension that recognized the historical slipperiness and dynamic movement of the "technology of race," Just as the relationship between Chang's scholarship and the contemporary telecommunication network of the internet further enframe my invocation of Coleman's proposition for an ethical encounter with the other – an anti-Colonial mediality – I believe it is important to seek to realign and reorient the overall consideration of the history of infrastructure of telecommunication and transportation not only as a history of violence, but one of resistance as well. This reorientation must as its first order of business insist on the presence, worth, and humanity of the racialized.

These modes of resistance and extension of self towards an ethical encounter with others have continued in the realm of content moderation. Since 2020, Foxglove, a non-profit advocacy group comprised of "lawyers, technology experts and communications specialists," has released

a series of open letters to Facebook and its subcontracting entities like Accenture and Covalen on behalf of hundreds of content moderators to demand fair wages, better working conditions, and an "end to outsourcing" (*Open Letter from Content Moderators*, 2020). In Ireland, another content moderator site away from the American centre, there are efforts now to push for legislative change to end outsourcing and for building stronger protection for the employees working on behalf of the tech companies to do their most violent and harmful work on the platform. In October of 2021, Sarah Plunkett became one of the first content moderators to publicly give evidence to a parliamentary committee of her working conditions. Working for the Dublin-based outsourcing firm Covalen, Plunkett worked on Facebook's platform, training its AI model to spot and block the platform's "most abhorrent content – including graphic child abuse and violence" (Nast, 2021). These efforts have sought to place a "duty of care" on the companies that do the hiring, whether they be Facebook or BPO companies. These movements for political pressure have also coincided with over "30 workers [bringing] personal injury cases against Facebook and third-party contractors across Europe claiming they have suffered from PTSD" (Nast, 2021). Indeed there has been a wave of legal challenges lodged by content moderators from across Europe, America, and Africa against social media platforms. In January of 2022, a Barcelona-based former Facebook content moderator filed preliminary proceedings at the Irish High Court in Dublin, against both Facebook and CCC Barcelona Digital, which is owned by Telus, "the global call centre giant" (Burke, 2022), and in May of 2022, Daniel Motaung, a former content moderator and union organizer at Sama, filed a lawsuit against "Facebook's parent company Meta and Samasource Kenya EPZ Ltd., Meta's content review subcontractor in the East African country" (Benson, 2022).

These kinds of political and legal scrutiny have had some positive impacts on the way subcontracting firms operate, with Accenture now asking its content moderators in the United States and Europe to sign a waiver acknowledging that they may develop PTSD and other mental health disorders prior to starting their work (Perrigo, 2022a). Following a *TIME* investigation that exposed the horrible working conditions in Kenya, Sama, a sub-contracting firm that touts its "integrated ML-powered platform with human-in-the-loop validation," has increased the salary of its content moderators between 30 to 50 percent. In 2020, Facebook paid out $52 million to fund mental health treatment for some of its American content moderators following a lawsuit centered on mental health problems stemming from their work, including PTSD (Perrigo, 2022a). The lawsuit involved more than 10,000 former and current content moderator in four U.S. states (Elliot & Parmar, 2020), and marked a crucial precedent in tasking social media companies like Facebook with tangible reforms to their business strategy and platform design.

But at the same time, whether these immediate responses to this scrutiny are enough is far from clear. Sama's content moderators, for example, have shared mixed feelings about their raise, noting that, while good news, a such raise is ultimately not enough: "Even with the pay rise, Sama employees remain some of Facebook's lowest-paid workers anywhere in the world" (Perrigo, 2022b). In the case of the $52 million lawsuit, its class representation ultimately left out all of the global south workers from India and the Philippines, where legal frameworks for labour protection do not exist to a same degree of accountability for the US firms operating there, a legal reality that is in part the design of those seeking to attract international capital and the international capital holders that seek to benefit from such an arrangement (Elliot & Parmar, 2020). Amidst this complex picture, asymmetry of legal protection is an additional area of

concern that must be investigated while continuing to push for worker solidarity and platform accountability.

What next?

As digital spaces expand and new terrains open up, how content moderation will operate and how those new spaces will be governed are becoming central issues of contention and debate. The "novel" conversations around the metaverse already include the urgent sounding refrain of "who will moderate these spaces/platforms and how?" as well as the counter-refrain: "[When talking about metaverse, while Meta may seek an "almost Disney level" of safety], moderating how users speak and behave at any meaningful scale is practically impossible" (Murphy, 2021). This concern is in part already bearing out with beta testers of Meta's VR social media platform, Horizon Worlds, reporting that they have been sexually harassed (Basu, 2021). Along with digital contents and artefacts, users of these new domains of social networking are being put in potential harm's way of other users' behaviours and bodily gestures, things that lack a categorically neat measurement or shape against which moderation guideline can be rote-applied by machine. No clear understanding of how content moderation will be implemented has yet been offered. This in part seems to presage a future media environment that will require further saturation of content moderators who will in turn be tasked with real-time monitoring and processing of vast amounts of incidence clips that metaverse users report.

In anticipation of such future platforms, the engineers and executives of Meta are already looking to intensify their moderation efforts. Andrew Bosworth, Meta's latest Chief Technology Officer (CTO), has been on record as saying that Meta's content moderation will lean on existing community rules, while having a "stronger bias towards enforcement along some sort of spectrum of warning, successively longer suspensions, and ultimately expulsion from multi-user

spaces" (Murphy, 2021). With "toxic behaviours" and bullying having the potential to be exacerbated in the "immersive nature of virtual reality," a finding that Facebook's Oculus Division made in their 2019 study, content moderators of the metaverse seem primed for even more exposure to harm as the demand for moderation to be real-time, immediate, and present becomes foundational to this new domain of social networking (Blackwell et al., 2019).

Based on my research so far, the solution does not seem to lie in simply increasing the number of workers or otherwise intensifying the existing form of content moderation, nor can it be found in a hopeful future where artificial intelligence will take on our dirty work. Meanwhile, content moderation continues to be a necessary function and labour the smooth running social media networks, and the human element of this process seems inevitable in our envisioning of future social media environments as well. The work that is done in the domain of content moderation is crucial in fostering a sense of community through safe connections and in preventing harm from reaching the vulnerable that seek safe haven in digital communities.

Hence the necessary question of who should be doing this work and under what circumstances. How can we create a situation where the work of content moderation can be engaged uncoerced and with fair compensation? How could we organize this labour so that their redistribution no longer simply replicates the historical continuity of the infrastructures that I explored in this thesis? Our questioning must also include probes into the fundamental logics and assumptions about the configuration of content moderation labour. Should the labour of content moderation be done on a commercial scale, and if so, what would that mean for the future? Can we imagine content moderation that is not undertaken by a racialized labour force, and what would that mean for compensation structures? Is it possible to adequately and fully gauge the risk of the work and provide commensurate compensation, including that of hazard pay? Would

it be right or ethical to continue to put certain classes of people in harm's way regardless of their working condition or compensation? These questions do not have simple answers.

The Cleaners (2018) closes with footage of two different figures: one is Mark Zuckerberg at an annual conference for Facebook and the other is an unnamed content moderator working at their station. In the footage, Zuckerberg delivers a keynote speech to an audience of computer engineers and product managers and shares the message of unity and connection that Facebook's products and platforms supposedly embody. Zuckerberg appeals to their optimism: "think that you can change the world." This change, he says, is a process that is made "one connection at a time," persistently "day after day after day." There is a thunderous applause. When it fades, his triumphant speech is followed by the quiet drone of an office space, where we see a moderator seated at their station. In this final dramatization of hidden labour, the moderator sits alone, facing a row of computer monitors as the sole source of light in the dark. The moderator screens the contents of Facebook, some that can be ignored, others that must be deleted. In the dark, these labourers help make the very connections that Zuckerberg imagines aloud. With Zuckerberg's hopeful words echoing, we realize that while it is difficult to straightforwardly answer who should be doing the work of moderating contents, these people are the ones that show up day after day after day to make the necessary decisions that help people connect, doing the work that is "more important than it's ever been before."

REFERENCES

Ackermann, K. (2020). *The Old Puppet Masters: Content Moderation on Computer Bulletin Board Systems*. Georgetown University.

Ahmad, S. (2019). *"It's Just the Job": Investigating the Influence of Culture in India's Commercial Content Moderation Industry* [Preprint]. SocArXiv. https://doi.org/10.31235/osf.io/hjcv2

Albarran, A. B. (Ed.). (2013). *The Social Media Industries*. Routledge.

Alhabash, S., & Ma, M. (2017). A Tale of Four Platforms: Motivations and Uses of Facebook, Twitter, Instagram, and Snapchat Among College Students? *Social Media + Society*, *3*(1), 205630511769154. https://doi.org/10.1177/2056305117691544

Arsht, A., & Etcovitch, D. (2018, March 2). *The Human Cost of Online Content Moderation*. Harvard Journal of Law & Technology. https://jolt.law.harvard.edu/digest/the-human-cost-of-online-content-moderation

Baginski, J., & Malecki, E. J. (2013). *The Oxford Handbook of the Digital Economy, edited by Martin Peitz and Joel Waldfogel* (Vol. 53). https://onlinelibrary.wiley.com/doi/10.1111/jors.12015_5

Barrett, P. M. (2020). *Who Moderates the Social Media Giants?: A A Call to End Outsourcing* (pp. 1–32). NYU Stern Centre for Business and Human Rights.

Basu, T. (2021, December 16). The metaverse has a groping problem already. *MIT Technology Review*. https://www.technologyreview.com/2021/12/16/1042516/the-metaverse-has-a-groping-problem/

Beals, M. (2021, September 21). *Facebook spent billions on safety, security over past 5 years* [Text]. TheHill. https://thehill.com/homenews/media/573213-facebook-spent-billions-on-safety-security-over-past-5-years

Benjamin, R. (2019). *Race after technology: Abolitionist tools for the new Jim code*. Polity.

Benson, E. A. (2022, June 3). *A South African ex. Employee of Facebook recently dragged the company to a Kenyan court; here's how the social media giant responded*. Business Insider Africa. https://africa.businessinsider.com/local/markets/a-south-african-former-employee-of-facebook-recently-sued-the-social-media-giant-in-a/mz9lny8

Biddle, S. (2021, October 12). Revealed: Facebook's Secret Blacklist of "Dangerous Individuals and Organizations." *The Intercept*. https://theintercept.com/2021/10/12/facebook-secret-blacklist-dangerous/

Blackwell, L., Ellison, N., Elliott-Deflo, N., & Schwartz, R. (2019). Harassment in Social Virtual Reality: Challenges for Platform Governance. *Proceedings of the ACM on Human-Computer Interaction*, *3*(CSCW), 1–25. https://doi.org/10.1145/3359202

Block, H., & Riesewieck, M. (Directors). (2018). *The Cleaners* [Documentary; Digital].

Bond, S. (2020, March 31). Facebook, YouTube Warn Of More Mistakes As Machines Replace Moderators. *NPR*. https://www.npr.org/2020/03/31/820174744/facebook-youtube-warn-of-more-mistakes-as-machines-replace-moderators

Brown, B. (2010). Materiality. In W. J. T. Mitchell & M. B. N. Hansen (Eds.), *Critical Terms for Media Studies*. The University of Chicago Press.

Burke, R. (2022, January 16). *Facebook moderators from across Europe bring first legal actions to Irish High Court*. Business Post. https://www.businesspost.ie/legal/facebook-moderators-from-across-europe-bring-first-legal-actions-to-irish-high-court-95b14b93

Canales, K. (2021a, October 25). *Facebook employees reportedly use internal farewell messages to criticize the company for not fixing its "net negative" influence on the world*. Business Insider. https://www.businessinsider.com/facebook-papers-employee-criticism-badge-posts-leaked-documents-whistleblower-2021-10

Canales, K. (2021b, October 29). *A departing Facebook employee said "omniscient robots owned by Mark Zuckerberg" won't save the company from its moderation problems*. Business Insider. https://www.businessinsider.com/facebook-papers-whistleblower-former-employee-said-ai-wont-save-company-2021-10

Carlson, C. R., & Cousineau, L. S. (2020). Are You Sure You Want to View This Community? Exploring the Ethics of Reddit's Quarantine Practice. *Journal of Media Ethics*, *35*(4), 202–213. https://doi.org/10.1080/23736992.2020.1819285

Carmi, E. (2019). *The Hidden Listeners: Regulating the Line from Telephone Operators to Content Moderators*. 19.

Chan, A., & McIntosh, A. (2022). Chinese Head Tax in Canada. In *The Canadian Encyclopedia*. https://www.thecanadianencyclopedia.ca/en/article/chinese-head-tax-in-canada

Chancellor, S., Pater, J. A., Clear, T., Gilbert, E., & De Choudhury, M. (2016). #thyghgapp: Instagram Content Moderation and Lexical Variation in Pro-Eating Disorder Communities. *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, 1201–1213. https://doi.org/10.1145/2818048.2819963

Chang, G. H. (2019). Chinese Railroad Workers and the US Transcontinental Railroad in Global Perspective. In G. H. Chang & S. F. Fishkin (Eds.), *The Chinese and the Iron Road: Buliding the Transcontinental Railroad*. Stanford University Press.

Chang, G. H., & Fishkin, S. F. (Eds.). (2019). *The Chinese and the Iron Road: Buliding the Transcontinental Railroad*. Stanford University Press.

Chang, G. H., Fishkin, S. F., & Obenzinger, H. (2019). Introduction. In G. H. Chang & Shelley Fisher Fishkin (Eds.), *The Chinese and the Iron Road: Buliding the Transcontinental Railroad*. Stanford University Press.

Chen, A. (2014, October 23). The Laborers Who Keep Dick Pics and Beheadings Out of Your Facebook Feed. *Wired*. https://www.wired.com/2014/10/content-moderation/

Chen, A., & Cassidy, C. (Directors). (2017, April 14). *The Moderators* [Documentary; Digital]. WIRED. https://www.wired.com/2017/04/watch-people-learn-filter-awfulness-dating-sites/

Chen, T. M., & Wang, V. (2010). Web Filtering and Censoring. *Computer*, *43*(3), 94–97. https://doi.org/10.1109/MC.2010.84

Chinese Exclusion Act (1882). (2021, September 8). National Archives. https://www.archives.gov/milestone-documents/chinese-exclusion-act

Chugg, J. (2021). *Amendment No. 3 to FORM S-1 REGISTRATION STATEMENT UNDER THE SECURITIES ACT OF 1933, TaskUs, Inc.* (pp. 1–278) [SEC Filing, Amendment to No. 3 to Form S-1]. United States Securities and Exchange Commission.

Chun, W. (2018). *Queerying Homiphily, in Pattern Discimination,*. 59–98.

Chun, W. H.-K. (2006). *Control and Freedom: Power and paranoia in the age of fiber optics*. MIT Press.

Coleman, B. (2009). Race as Technology. *Camera Obscura: Feminism, Culture, and Media Studies*, *24*(1), 177–207. https://doi.org/10.1215/02705346-2008-018

Complex, adj. (n.d.). In *OED Online*. Oxford University Press. Retrieved March 4, 2022, from http://www.oed.com/view/Entry/37672

C-SPAN (Director). (2021, October 5). *Facebook Whistleblower Frances Haugen testifies before Senate Commerce Committee*. https://www.youtube.com/watch?v=GOnpVQnv5Cw

Dearinger, R. (2016). *The filth of progress: Immigrants, Americans, and the building of canals and railroads in the West* (First edition). University of California Press.

Dixon, K. J. (2019). Landscapes of Change: Culture, Nature, and the Archaeological Heritage of the Transcontinental Railways in the North American West. In S. F. Fishkin & G. H. Chang (Eds.), *The Chinese and the Iron Road: Buliding the Transcontinental Railroad*. Stanford University Press.

Dwoskin, E. (2019, May 8). Inside Facebook, the second-class workers who do the hardest job are waging a quiet battle. *Washington Post*. https://www.washingtonpost.com/technology/2019/05/08/inside-facebook-second-class-workers-who-do-hardest-job-are-waging-quiet-battle/

Einwiller, S. A., & Kim, S. (2020). How Online Content Providers Moderate User-Generated Content to Prevent Harmful Online Communication: An Analysis of Policies and Their Implementation. *Policy & Internet*, *12*(2), 184–206. https://doi.org/10.1002/poi3.239

Elliot, V., & Parmar, T. (2020, July 22). *"The despair and darkness of people will get to you."* Rest of World. https://restofworld.org/2020/facebook-international-content-moderators/

Elmer, G. (2004). *Profiling machines: Mapping the personal information economy*. MIT Press.

Elmer, G. (2019). Prospecting Facebook: The limits of the economy of attention. *Media, Culture & Society*, *41*(3), 332–346. https://doi.org/10.1177/0163443718813467

Emerson, S. (2020, July 28). 'A Permanent Nightmare': Pinterest Moderators Fight to Keep Horrifying Content Off the Platform. *OneZero*. https://onezero.medium.com/a-permanent-nightmare-pinterest-moderators-fight-to-keep-horrifying-content-off-the-platform-4d8e7ec822fe

Enns, A. (2015). Introduction: The Media Philosophy of Sybille Krämer. In S. Krämer, *Medium, Messenger, Transmission: An Approach to Media Philosophy*. Amsterdam University Press. https://doi.org/10.1515/9789048524990

Facebook. (2020, December 18). Keeping People Safe and Informed About the Coronavirus. *Meta*. https://about.fb.com/news/2020/12/coronavirus/

Facebook. (2021, May 12). *Advertising Policies Related to Coronavirus (COVID-19)*. Facebook Business Help Center. https://www.facebook.com/business/help/1123969894625935

Fagan, F. (2020). Optimal social media content moderation and platform immunities. *European Journal of Law and Economics*, *50*(3), 437–449. https://doi.org/10.1007/s10657-020-09653-7

Fishkin, S. F. (2019). The Chinese as Railroad Builders after Promontory. In S. F. Fishkin & G. H. Chang (Eds.), *The Chinese and the Iron Road: Buliding the Transcontinental Railroad*. Stanford University Press.

Foxglove. (2021, May 13). The Committee on Enterprise, Trade and Employment heard directly from Facebook content moderator Isabella. Here's a recap of some of what she said 👇 https://t.co/Ge5KA0XhDP [Tweet]. *@Foxglovelegal*. https://twitter.com/Foxglovelegal/status/1392796853708730369

Galloway, A. R. (2010). Networks. In W. J. T. Mitchell & M. B. N. Hansen (Eds.), *Critical Terms for Media Studies*. The University of Chicago Press.

Ganesh, B., & Bright, J. (2020). Countering Extremists on Social Media: Challenges for Strategic Communication and Content Moderation. *Policy & Internet*, *12*(1), 6–19. https://doi.org/10.1002/poi3.236

Gilbert, D. (2019, December 3). Bestiality, Stabbings, and Child Porn: Why Facebook Moderators Are Suing the Company for Trauma. *Vice*. https://www.vice.com/en/article/a35xk5/facebook-moderators-are-suing-for-trauma-ptsd

Gilbert, D. (2021, May 12). *Facebook Is Ignoring Moderators' Trauma: 'They Suggest Karaoke and Painting.'* Vice. https://www.vice.com/en/article/m7eva4/traumatized-facebook-moderators-told-to-suck-it-up-and-try-karaoke

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

Gillespie, T. (2020). Content moderation, AI, and the question of scale. *Big Data & Society*, *7*(2), 205395172094323. https://doi.org/10.1177/2053951720943234

Government of Canada, S. C. (2019, May 3). *Measuring digital economic activities in Canada: Initial estimates*. https://www150.statcan.gc.ca/n1/pub/13-605-x/2019001/article/00002-eng.htm

Gow, W. (2019). The Chinese Railroad Worker in United States History Textbooks: A Historical Genealogy, 1849-1965. In S. F. Fishkin & G. H. Chang (Eds.), *The Chinese and the Iron Road: Buliding the Transcontinental Railroad*. Stanford University Press.

Gray, C. (2019, December 9). As a Facebook moderator I saw the worst of humanity. We need to be valued. *The Guardian*. https://www.theguardian.com/commentisfree/2019/dec/09/facebook-moderator-worst-of-humanity-valued

Greis, M., Alt, F., Henze, N., & Memarovic, N. (2014). I can wait a minute: Uncovering the optimal delay time for pre-moderated user-generated content on public displays. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1435–1438. https://doi.org/10.1145/2556288.2557186

Guevara, R. L. (2021, March 1). *In Content Moderation, Employees' Mental Health Comes First*. TaskUs. https://www.taskus.com/insights/best-practices-for-content-moderation-employee-psychological-health-and-the-critical-role-of-evidence-based-resiliency-programming/

Hern, A. (2019, December 4). Ex-Facebook worker claims disturbing content led to PTSD. *The Guardian*. https://www.theguardian.com/technology/2019/dec/04/ex-facebook-worker-claims-disturbing-content-led-to-ptsd

Hever, P., & Crowley, D. (2008). Introduction. In H. Innis, *The Bias of Communication* (2nd ed). University of Toronto Press.

Hsinya, H. (2019). Tracking Memory: Encounters between Chinese Railraod Workers and Native Americans. In G. H. Chang & S. F. Fishkin (Eds.), *The Chinese and the Iron Road: Buliding the Transcontinental Railroad*. Stanford University Press.

Innis, H. (1923). A History of the Canadian Pacific Railway. *The ANNALS of the American Academy of Political and Social Science*, *108*(1), 228–229. https://doi.org/10.1177/000271622310800148

Innis, H. (2008). *The Bias of Communication* (2nd ed). University of Toronto Press.

Jardine, E. (2019). Online content moderation and the Dark Web: Policy responses to radicalizing hate speech and malicious content on the Darknet. *First Monday*. https://doi.org/10.5210/fm.v24i12.10266

Jereza, R. (2021). Corporeal moderation: Digital labour as affective good. *Social Anthropology*, *29*(4), 928–943. https://doi.org/10.1111/1469-8676.13106

Jhaver, S., Birman, I., Gilbert, E., & Bruckman, A. (2019). Human-Machine Collaboration for Content Regulation: The Case of Reddit Automoderator. *ACM Transactions on Computer-Human Interaction*, *26*(5), 1–35. https://doi.org/10.1145/3338243

Kapoor, K. K., Tamilmani, K., Rana, N. P., Patil, P., Dwivedi, Y. K., & Nerur, S. (2018). Advances in Social Media Research: Past, Present and Future. *Information Systems Frontiers*, *20*(3), 531–558. https://doi.org/10.1007/s10796-017-9810-y

Kaye, D. (2018). *Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression*. United Nations Human Rights Office of the High Comissioner. https://www.undocs.org/A/HRC/38/35

Krämer, S. (2015). *Medium, Messenger, Transmission: An Approach to Media Philosophy*. Amsterdam University Press. https://doi.org/10.1515/9789048524990

Lampe, C., & Resnick, P. (2004). Slash(dot) and burn: Distributed moderation in a large online conversation space. *Proceedings of the 2004 Conference on Human Factors in Computing Systems - CHI '04*, 543–550. https://doi.org/10.1145/985692.985761

Lavallé, O., & Marshall, T. (2022). Canadian Pacific Railway. In *The Canadian Encyclopedia*. https://www.thecanadianencyclopedia.ca/en/article/canadian-pacific-railway

Link, D., Ling, J., Hoffjann, J., & Hellingrath, B. (2016). A Semi-Automated Content Moderation Workflow for Humanitarian Situation Assessments: *International Journal of Information Systems for Crisis Response and Management*, *8*(2), 31–49. https://doi.org/10.4018/IJISCRAM.2016040103

Manovich, L. (2001). *The Language of New Media* (8. print). MIT Press.

Marte-Wood, A. S., & Santos, S. D. (2021). Circuits of Care: Filipino Content Moderation and American Infostructures of Feeling. *Verge: Studies in Global Asias*, *7*(2), 101. https://doi.org/10.5749/vergstudglobasia.7.2.0101

Matsakis, L., & Martineau, P. (2020, March 10). Coronavirus Disrupts Social Media's First Line of Defense. *Wired*. https://www.wired.com/story/coronavirus-social-media-automated-content-moderation/

Meta. (2021a, June 11). *Helping reviewers make the right calls*. Transparency Center. https://transparency.fb.com/enforcement/detecting-violations/making-the-right-calls/

Meta. (2021b, June 11). *The people behind Facebook's review teams*. Transparency Center. https://transparency.fb.com/enforcement/detecting-violations/people-behind-our-review-teams/

Meta. (2021c, September 21). Our Progress Addressing Challenges and Innovating Responsibly. *Meta*. https://about.fb.com/news/2021/09/our-progress-addressing-challenges-and-innovating-responsibly/

Meta. (2022a, January 19). *How technology detects violations*. Transparency Center. https://transparency.fb.com/enforcement/detecting-violations/technology-detects-violations/

Meta. (2022b, January 19). *How technology helps prioritize review*. Transparency Center. https://transparency.fb.com/enforcement/detecting-violations/technology-helps-prioritize-review/

MicroSourcing. (n.d.-a). *Find Out More About Our History*. Retrieved April 17, 2022, from https://www.microsourcing.com/about-us/history/

MicroSourcing. (n.d.-b). *Offshore Staffing*. Retrieved May 28, 2022, from https://www.microsourcing.com/how-it-works/offshore-staffing/

MicroSourcing. (n.d.-c). *Why Outsource To The Philippines?* MicroSourcing. Retrieved December 14, 2021, from https://www.microsourcing.com/why-offshore/why-the-philippines/

MicroSourcing. (2021). *Job Profile: Content Moderator Outsourcing Philippines*. https://www.microsourcing.com/build-your-team/job-roles/content-moderator/

Ministry of International Trade. (n.d.). *Building the Railway*. Province of British Columbia; Province of British Columbia. Retrieved March 15, 2022, from https://www2.gov.bc.ca/gov/content/governments/multiculturalism-anti-racism/chinese-legacy-bc/history/building-the-railway

Momeni, E. (2012). Semi-automatic semantic moderation of web annotations. *Proceedings of the 21st International Conference on World Wide Web*, 167–172. https://doi.org/10.1145/2187980.2188003

Morozov, E. (2013). *To save everything, click here: The folly of technological solutionism* (First edition). PublicAffairs.

Mulla, S., & Palave, A. (2016). Moderation technique for sexually explicit content. *2016 International Conference on Automatic Control and Dynamic Optimization Techniques (ICACDOT)*, 56–60. https://doi.org/10.1109/ICACDOT.2016.7877551

Murphy, H. (2021, November 12). How will Facebook keep its metaverse safe for users? *Financial Times*. https://www.ft.com/content/d72145b7-5e44-446a-819c-51d67c5471cf

Myers West, S. (2018). Censored, suspended, shadowbanned: User interpretations of content moderation on social media platforms. *New Media & Society*, *20*(11), 4366–4383. https://doi.org/10.1177/1461444818773059

Nast, C. (2021, November 6). Facebook’s content moderators are fighting back. *Wired UK*. https://www.wired.co.uk/article/facebook-content-moderators-ireland

Newton, C. (2019, February 25). *The Trauma Floor: The secret lives of Facebook moderators in America*. The Verge. https://www.theverge.com/2019/2/25/18229714/cognizant-facebook-content-moderator-interviews-trauma-working-conditions-arizona

Newton, C. (2020, May 12). *Facebook will pay $52 million in settlement with moderators who developed PTSD on the job*. The Verge. https://www.theverge.com/2020/5/12/21255870/facebook-content-moderator-settlement-scola-ptsd-mental-health

Nuttman-Shwartz, O., & Green, O. (2021). Resilience truths: Trauma resilience workers' points of view toward resilience in continuous traumatic situations. *International Journal of Stress Management*, *28*(1), 1–10. https://doi.org/10.1037/str0000223

O'Connell, J. (2019, March 30). Facebook's dirty work in Ireland: 'I had to watch footage of a person being beaten to death.' *The Irish Times*. https://www.irishtimes.com/culture/tv-radio-web/facebook-s-dirty-work-in-ireland-i-had-to-watch-footage-of-a-person-being-beaten-to-death-1.3841743

O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy* (First edition). Crown.

Open letter from content moderators re: Pandemic. (2020, November 18). Foxglove. https://www.foxglove.org.uk/news/open-letter-from-content-moderators-re-pandemic

Oyedemi, T. D. (2019). Global Digital Capitalism: Mark Zuckerberg in Lagos and the Political Economy of Facebook in Africa. *International Journal of Communication*, *13*, 2045–2061.

Oyedemi, T. D. (2020). Digital coloniality and 'Next Billion Users': The political economy of Google Station in Nigeria. *Information, Communication & Society*, 1–15. https://doi.org/10.1080/1369118X.2020.1804982

Parks, L. (2015). "Stuff You Can Kick": Toward a Theory of Media Infrastructures. In P. Svensson & D. T. Goldberg (Eds.), *Between humanities and the digital* (pp. 355–373). The MIT Press.

Perrigo, B. (2022a, February 14). *Inside Facebook's African Sweatshop*. Time. https://time.com/6147458/facebook-africa-content-moderation-employee-treatment/

Perrigo, B. (2022b, March 2). *Facebook Content Moderators in Kenya Are Receiving a Pay Rise Following TIME Investigation*. Time. https://time.com/6153778/facebook-moderators-kenya-sama/

Prashad, V. (2012). *The Poorer Nations: A possible history of the Global South*. Verso.

R.A. 7925. (1995, March 1). https://lawphil.net/statutes/repacts/ra1995/ra_7925_1995.html

Reuters Staff. (2007, October 16). UPDATE 1-Facebook settles New York child safety probe. *Reuters*. https://www.reuters.com/article/facebook-settlement-idUSN1617823820071016

Richardson, H. C. (2004). *The Death of Reconstruction: Race, labor, and politics in the post-Civil War North, 1865 - 1901* (1. Harvard Univ. Press paperback ed). Harvard University Press.

Rigaud, S. (Director). (2019). *Ghost Workers*. Java Films. https://torontopl.kanopy.com/video/ghost-workers

Roberts, S. T. (2014). *Behind the Screen: The Hidden Digital Labor of Commercial Content Moderation* [Dissertation]. University of Illinois at Urbana-Champaign.

Roberts, S. T. (2016). Digital Refuse: Canadian Garbage, Commercial Content Moderation and the Global Circulation of Social Media's Waste. *Media Studies Publications*, *14*, 12.

Roberts, S. T. (2017). *Content moderation*. https://escholarship.org/uc/item/7371c1hf

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

Roberts, S. T. (2020, April 8). The Great A.I. Beta Test. *Slate*. https://slate.com/technology/2020/04/coronavirus-facebook-content-moderation-automated.html

Roh, D. S., Huang, B., & Niu, G. A. (Eds.). (2015). *Techno-Orientalism: Imagining Asia in speculative fiction, history, and media*. Rutgers University Press.

Roth, E. (2021, October 19). *Facebook content moderators protest low wages with mobile billboard*. The Verge. https://www.theverge.com/2021/10/19/22726915/facebook-content-moderators-protest-low-wages

Sajed, A. (2020, July 27). From the Third World to the Global South. *E-International Relations*. https://www.e-ir.info/2020/07/27/from-the-third-world-to-the-global-south/

Santa Clara Law. (n.d.). *Content Moderation & Removal at Scale*. Santa Clara Law. Retrieved July 16, 2021, from https://law.scu.edu/event/content-moderation-removal-at-scale/

Satariano, A., & Isaac, M. (2021a, August 31). The Silent Partner Cleaning Up Facebook for $500 Million a Year. *The New York Times*. https://www.nytimes.com/2021/08/31/technology/facebook-accenture-content-moderation.html

Satariano, A., & Isaac, M. (2021b, September 1). *How Accenture makes millions scrubbing Facebook's sordid content* [The Economic Times]. https://economictimes.indiatimes.com/tech/information-tech/accenture-the-silent-partner-cleaning-up-facebook-for-500-million-a-year/articleshow/85828288.cms

Scott, M., & Kayali, L. (2020, October 21). *What happened when humans stopped managing social media content*. POLITICO. https://www.politico.eu/article/facebook-content-moderation-automation/

Shaw, M. J. (2006). *E-Commerce and the Digital Economy* (Vol. 4). Routledge, Taylor & Francis Group.

Siapera, E., & Viejo-Otero, P. (2021). Governing Hate: Facebook and Digital Racism. *Television & New Media*, *22*(2), 112–130. https://doi.org/10.1177/1527476420982232

Silver, E. (2018, July 26). Hard Questions: Who Reviews Objectionable Content on Facebook — And Is the Company Doing Enough to Support Them? *Meta*. https://about.fb.com/news/2018/07/hard-questions-content-reviewers/

Simplr. (2019, January 10). *TaskUs Pricing*. Simplr. https://www.simplr.ai/learn/taskus-pricing

Slack, J. D., & Wise, J. M. (2015). *Culture and Technology: A Primer* (Second edition). Peter Lang.

Spillers, H. J. (1987). Mama's Baby, Papa's Maybe: An American Grammar Book. *Diacritics*, *17*(2), 64. https://doi.org/10.2307/464747

Stanford, L. (1865). *Statement Made to the President of the United States, and Secretary of the Interior, of the Progress of the Work*. H.S. Crocker & Co., Printers.

Starosielski, N. (2015). *The Undersea Network*. Duke University Press.

Statista. (n.d.). *Social Media & User-Generated Content*. Statista. Retrieved November 27, 2021, from https://www.statista.com/markets/424/topic/540/social-media-user-generated-content//

Statista Research Department. (2021a, February 25). *Social media—Statistics & Facts*. Statista. https://www-statista-com.ezproxy.library.yorku.ca/topics/1164/social-networks/

Statista Research Department. (2021b, November 16). *Most used social media 2021*. Statista. https://www-statista-com.ezproxy.library.yorku.ca/statistics/272014/global-social-networks-ranked-by-number-of-users/

Statista Research Department. (2022a, February 8). *Facebook: Revenue and net income*. Statista. https://www.statista.com/statistics/277229/facebooks-annual-revenue-and-net-income/

Statista Research Department. (2022b, April 28). *Facebook MAU worldwide 2022*. Statista. https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/

Steer, J. (2017, April 17). Facebook describes timeline of events in Cleveland murder, announces review. *Fox 8 Cleveland WJW*. https://fox8.com/news/facebook-describes-timeline-of-events-in-cleveland-murder-announces-review/

Steiger, M. (2020). *Building a Resilient Workforce: Programming for Commercial Content Moderation Staff* [St. Mary's University]. https://www.proquest.com/openview/0a8fd4bb3d328027ea84f1590fc9c175/1?pq-origsite=gscholar&cbl=18750&diss=y

Steiger, M., Bharucha, T. J., Venkatagiri, S., Riedl, M. J., & Lease, M. (2021). The Psychological Well-Being of Content Moderators: The Emotional Labor of Commercial Moderation and Avenues for Improving Support. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–14. https://doi.org/10.1145/3411764.3445092

Taffel, S. (2018). Hopeful Extinctions? Tesla, Technological Solutionism and the Anthropocene. *Culture Unbound*, *10*(2), 163–184. https://doi.org/10.3384/cu.2000.1525.2018102163

TaskUs. (n.d.). *The Human Power Behind AI & ML*. TaskUs. Retrieved May 27, 2022, from https://www.taskus.com/ai-services/

Terranova, T. (2000). Free Labor: Producing Culture for the Digital Economy. *Social Text*, *18*(2 (63)), 33–58. https://doi.org/10.1215/01642472-18-2_63-33

The Chinaman as a Railroad Builder. (1869). *Scientific American*, *XXI*(5), 75.

The Facebook Company Is Now Meta. (2021, October 28). *Meta*. https://about.fb.com/news/2021/10/facebook-company-is-now-meta/

The Facebook Files. (2021, October 1). *Wall Street Journal*. https://www.wsj.com/articles/the-facebook-files-11631713039

Towns, A. R. (2019). (Dark) Pure War: Virilio, the Cinematic, and the Racial. *Media Theory*, *3*(2), 1–16.

Towns, A. R. (2020). Toward a Black media philosophy. *Cultural Studies*, *34*(6), 851–873. https://doi.org/10.1080/09502386.2020.1792524

Veglis, A. (2014). Moderation Techniques for Social Media Content. In G. Meiselwitz (Ed.), *Social Computing and Social Media* (pp. 137–148). Springer International Publishing. https://doi.org/10.1007/978-3-319-07632-4_13

Wagner, K. (2017, May 3). *Facebook is hiring another 3,000 people to pull down violent and inappropriate content*. Vox. https://www.vox.com/2017/5/3/15531478/facebook-hiring-3000-people-violent-inappropriate-video-content-post

Watson, A. J. (2008). Introduction to the Second Edition. In H. Innis, *The Bias of Communication* (2nd ed). University of Toronto Press.

Yung, J. (Ed.). (2006). *Chinese American voices: From the gold rush to the present*. Univ. of California Press.

Zuckerberg, M. (2017, May 3). *Facebook* [Facebook Post]. Facebook. https://www.facebook.com/zuck/posts/10103695315624661?notif_t=notify_me¬if_id=1493820261300939

Zuckerberg, M. (2022, March 18). *Facebook, Inc. Press Call #8368103* [Press Call Transcript].

Zwass, V. (2006). Series Editor's Introduction. In V. Zwass (Ed.), *E-Commerce and the Digital Economy* (Vol. 4). Routledge, Taylor & Francis Group.