**Human Rights Advocacy in Silicon Valley: Civil Society's Perceptions of Platform
Accountability Efforts and Stakeholder Engagement Processes**

by

Kayla L. Brown

A thesis submitted in satisfaction of
the honors designation in Media Studies

University of California, Berkeley

Spring 2020

Advisor: Meeta Rani Jha, PhD

Acknowledgements

A special thank you to my thesis advisor, Meeta Rani Jha, PhD, as well the entire Media Studies department for all of the support and guidance throughout my undergraduate years. I would also like to thank Berkeley Law's Human Rights Center and the Center for Responsible Business at Berkeley Haas. I credit a large amount of my personal, professional, and academic growth to the coursework I've taken and the mentorship and research experience I've received through both of these centers.

Table of Contents

Section I: Introduction

Background

Relevance

Thesis Overview

Section II: Literature Review

Defining *Social Media*, *Platform*, and *Platform Company*

Social Media Platforms and Democratic Discourse

Social Media and Human Rights Advocacy

Platforms and Corporate Interests

Legal and Policy Landscape

Governance by Platform Companies

Defining *Platform Accountability* and *Content Governance*

Platform Accountability Efforts via Multistakeholderism

Civil Society, Corporate Accountability, and Social Change

Section III: Methodology

Theoretical Frames

Research Questions

Research Methods

Data Collection

Scope and Limitations

Section IV: Findings

Question 1

Question 2

Question 3

Section V: Discussion & Conclusion

Relevance

Further Research

Conclusion

Section I: Introduction

Background

On May 6, 2020, Facebook announced the first members of its new Oversight Board, nearly two and a half years after CEO Mark Zuckerberg first announced its concept in late 2018. The Board, analogous to Facebook's "Supreme Court" for content moderation and enforcement decisions, is an independent body that was established by Facebook and will make binding decisions about content posted to Facebook and Instagram, in addition to providing recommendations on the company's content policy ("Establishing Structure and Governance for an Independent Oversight Board," 2019). The Board notes that its decisions will be based on respect of freedom of expression and human rights ("Oversight Board," 2020).

While the Board is an unprecedented body in the realm of content governance, its establishment does not come entirely as a surprise. Over the years, Facebook has received widespread backlash and criticism for the negative harms associated with its products and services and for overlooking human rights considerations. Facebook, in particular, has perhaps captured the most negative public attention of any social media platform, and in many ways, the company has become the colloquial surrogate used interchangeably with social media at large. However, Facebook is certainly not the only major social media platform that has been engulfed in scandal.

A growing acknowledgement of the cultural and economic power that social media platform companies possess has also been met with a call for these companies to be held accountable for their negative social impacts. Civil society organizations in particular have been instrumental in drawing attention to the human rights impacts and considerations associated with

social media platforms. Increasingly, civil society organizations are also working directly with platform companies to improve their products, services, and governance processes.

Relevance

Facebook's Oversight Board can be situated within what John Bowers and Jonathan Zittrain refer to as the new "Process" era of internet governance, in which companies are beginning to focus on building out content governance *processes* to improve democratic transparency and legitimacy (Bowers & Zittrain, 2020). These processes are being built out, in part, due to the absence of more formal governance mechanisms. As transnational corporations, platform companies yield an enormous amount of economic, cultural and political power. However, absent any formal governance mechanism to hold these platforms accountable to their harms and ensure that the rights of the users they serve are being protected, civil society organizations fill an important gap and can act as a counterweight to both state and corporate power. The role that civil society organizations play in this new "Process" era of content governance and platform accountability is therefore worthy of examination. Human rights organizations, in particular, serve a critical role. They are advocates for society's most vulnerable populations, many of whom are most affected when platform companies overlook their human rights obligations. Furthermore, human rights norms and international human rights law can lend a sense of legitimacy to the privately written user policies and guidelines that inevitably impact whether and how we can express ourselves on social media platforms.

Thesis Overview

This research project evaluates civil society's perceptions of human rights advocacy and stakeholder engagement within the particular context of content governance and platform

accountability. It is informed by critical theories of new media and scholarship on social movements and social change, and it seeks to answer the following three research questions:

- **Question 1:** What *strategies* do civil society organizations use in order to advocate for human rights within the context of content governance and platform accountability?
- **Question 2:** How do these organizations *understand* their role in ongoing human rights advocacy and platform accountability efforts, relative to platform companies?
- **Question 3:** What major *challenges* and *opportunities* do human rights experts and civil society organizations identify when it comes to holding platform companies accountable to their human rights impacts?

Findings reveal that, within the context of platform accountability and content governance, organized civil society advocates for human rights by applying external pressure through media coverage and by networking directly with platform companies, many of which are hiring dedicated human rights leads and establishing more robust stakeholder engagement or governance processes; corporate interests underpin both strategies, however, the individuals interviewed identified the former as the most effective. Although organized civil society is an important counterweight to corporate power, findings reveal that business interests are still a significant barrier to enacting meaningful social change. Findings also suggest that good faith by and good actors within platform companies are ultimately not enough, reinforcing the important role organized civil society plays in increasing democratic accountability, even as private corporations begin to create processes to govern themselves.

Section II: Literature Review

Defining *Social Media*, *Platform*, and *Platform Company*

Colloquially, most people associate “social media” with a wide range of sites and applications. YouTube, Twitter, Facebook, Instagram, LinkedIn, Reddit, TikTok are some of the most popular platforms in the United States. In her book *The Culture of Connectivity: A Critical History of Social Media*, media studies scholar José van Dijck offers a rough definition of social media, which she describes as:

...a group of Internet-based applications that build on the ideological and technical foundations of Web 2.0 and that allow the creation and exchange of user-generated content, form a new online layer through which people organize their lives (van Dijck, 2013)

She notes that, in the early 2000’s, the promise of *connectedness* is what originally drove users to these platforms. Today, a handful of powerful companies dominate the global market and run the world’s most popular platforms, and “social media” is now synonymous with many of the aforementioned companies (van Dijck, 2013).

The term “platform” is typically used to describe the major social media companies, however, “platform” can also be used in reference to the services these companies offer. YouTube, for example, would be considered a platform upon which users share, view and comment on videos. In academia, the term “platform” has various denotations. Furthermore, as communication scholar Tarleton Gillespie notes, social media companies did not always frame themselves as platforms (Gillespie, 2018). It is therefore worth briefly exploring how the term “platform” has evolved throughout the internet’s history and since been used by private companies who seek to frame their products and services in specific and strategic ways.

Following the burst of the dot-com bubble, O'Reilly Media outlined a vision for the internet's next generation, or Web 2.0. O'Reilly described Web 2.0 as "a set of principles and practices" that would constitute the internet's new architecture. The first principle of Web 2.0 reimagined the "web as a platform" defined primarily by user contribution, collaboration, and participation (O'Reilly, 2007). However, as Matthew Allen notes in his discourse analysis, it is also useful to conceptualize Web 2.0 and as a "rhetorical technology" (Allen, 2013). In other words, the purpose of Web 2.0 discourse, in the wake of the dot-com bust, was to clarify the internet's contemporary role as a *computational* platform for new services, while also legitimizing future developments and reimaginings of the world wide web (Allen, 2013; Helmond, 2015). A computational platform is simply "a system that can be reprogrammed and therefore customized by outside developers" (Bogost & Montfort, 2009).

Social media companies have since strategically deployed the term "platform" when framing their products and services, drawing on the term's *computational*, *architectural*, *figurative*, and *political* connotations to situate themselves within existing Web 2.0 discourse and market themselves to users, advertisers, and shareholders. This framing also obscures the ways in which these companies function as traditional media gatekeepers and play a key role in shaping public discourse (Gillespie, 2010; Helmond, 2015). The term "platform" is not only embedded with layers of ideological value but also has become the colloquial default for how we conceptualize companies like Facebook (Srnicsek, 2016). However, as Gillespie notes, this use of the term "too easily equates a site with the company that offers it" (Gillespie, 2018). Social media companies should instead be thought of as "socio-technical assemblages and complex institutions" (Gillespie, 2017). Building upon Srnicsek and Gillespie's discursive work, Robert

Gorwa offers two working definitions. A *platform* refers to a specific online, data-driven app or service, while a *platform company* refers to the corporation that deploys that service (Gorwa, 2019a). For the purpose of clarity and consistency, this paper will rely on Gorwa's definition of *platform* and *platform company*, in addition to van Djick's definition of *social media*.

Social Media Platforms and Democratic Discourse

It is important to understand how platform companies frame their normative role in society in addition to how their products and services are actually deployed and situated. Discourse analysis reveals that platform companies have historically framed their products and services as inherently democratizing tools for self-expression and communication. Facebook CEO Mark Zuckerberg has argued that the company's platform enabled social movements like #BlackLivesMatter to spread in unprecedented ways, and he frequently cites Facebook's mission to "bring people closer together" (Romm, 2019). A discursive analysis of Mark Zuckerberg's public speeches spanning a period of ten years also reveals that the CEO has continuously emphasized language centering on sharing and connection (Hoffmann et al., 2018). Similarly, when protests broke out challenging long-standing authoritarian regimes in countries like Libya, Tunisia, and Egypt during the Arab Spring, these uprisings were quickly and commonly referred to as the "Twitter Revolutions," given that activists frequently took to Twitter's platform in order to organize and share events in real-time (Joseph, 2012).

Although research illustrates that platform use was instrumental in spreading information and increasing participation in protests during the Arab Spring, critics have also complicated the narrative of technological determinism asserting that platforms are democratic forces for good in and of themselves. In a series of surveys and interviews that sampled over 1,200 people who

partook in demonstrations in Tahrir, Egypt, researchers found that social media platforms played a “crucial role” in disseminating information, but were nevertheless part of a “complex intertwining of online and offline spheres” (Tufekci & Wilson, 2012). In other words, platforms do not determine social or political environments, but they are certainly a part of them (Morozov, 2009).

Clay Shirky argues that social media platforms have become “coordinating tools in nearly all of the world’s political movements,” but notes that these movements aren’t inherently or automatically democratic. They can also be regressive. (Shirky, 2011). Similarly, Rebecca MacKinnon argues that, although social media platforms can be used to coordinate and organize political movements, they can also be co-opted and deployed by governments to support authoritarian or oppressive regimes. She writes in her book *Consent of the Networked*: “though the technology used for coordinating and organizing may be politically neutral, the context in which it is deployed is rarely so” (MacKinnon, 2012). It is clear that platforms, with regards to their democratic potential, are ultimately inseparable from the larger social, political and economic networks in which they are used.

Social Media and Human Rights Advocacy

It is obvious that social media can be used in many different ways by many different actors. Given the scope of this research project, an overview of how social media platforms have been used within the particular context of human rights advocacy and human rights documentation is provided. As discussed with regards to democratic discourse, social media platforms were instrumental coordinating tools throughout the Middle East during the Arab Spring. Platforms, along with other information and communications technologies (ICTs), were

used to coordinate, document, and broadcast in-person events. (Shirky, 2011). Social media platforms have also become vital tools for human rights documentation, advocacy and legal accountability. Content posted to these platforms depicting human rights violations or evidence of war crimes can serve as collective memory and shed light on acts of violence against marginalized communities that may otherwise be undocumented, obscured, or erased (Abdul Rahman Al Jaloud et al., 2019). In the United States, videos posted and shared across social media platforms depicting acts of police brutality, particularly against black men, have contributed to movements like #BlackLivesMatter and led to greater calls for accountability and justice (Carter, 2016).

Outside of the United States, human rights investigators and civil society organizations have begun to use open-source information, particularly content posted to social media platforms, for human rights reporting and legal accountability (Dubberley et al., 2020). When Myanmar's security forces were accused of committing serious human rights violations against the Rohingya minority in Rakhine state in 2016 and 2017, Myanmar officials sealed off the region from outside observers such as journalists and the United Nations, making on-the-ground investigations virtually impossible. Fortunately, images and videos posted to social media, combined with satellite imagery and other open-source tools, allowed researchers and investigators to verify human rights violations and combat state misinformation (Dubberley et al., 2020).

Visual content depicting human rights violations have also been used to prosecute war criminals. In August 2017, the International Criminal Court (ICC) issued an arrest warrant for Mahmoud Mustafa Busayf Al-Werfalli, a Libyan military commander accused of committing or

ordering over thirty-three murders in Benghazi (Irving, 2017). This was the ICC's first arrest warrant based in part on evidence collected from social media platforms (Mehandru & Koenig, 2019b). Platforms may not have been built to document evidence of war crimes and human rights violations, however, they have undoubtedly become de-facto archives of this content.

Platforms and Corporate Interests

As discussed, platforms sit within larger socio-political contexts and can be instrumental tools for human rights advocacy and accountability. Platforms therefore do have democratic value, albeit inseparable from the larger social and political contexts in which they operate. However, platforms are also critically and practically inseparable from the large multinational corporations that own them. Understanding the platform economy and the business interests of platform companies is therefore a crucial component of platform accountability. Business interests complicate the narrative that platforms are either neutral or inherently democratizing. Platforms may increase access to information and conversation, but the economic interests of platform companies are nevertheless still embedded within the very structure of their products and services (Gillespie, 2017).

Born out of Web 2.0, social media platforms were founded upon ideological notions of connectedness, community, participation and collaboration—in other words, the “social” in social media (van Dijck, 2013). In comparison to traditional forms of mass media, platforms blurred the lines between media producer and media consumer, creating a new class of “prosumers” (Fuchs, 2015; van Dijck, 2013). As sites of social exchange, many platforms indeed started out as community initiatives. However, platforms are built and operated by companies, so the ways that platforms *organize* social exchange is, at least in part, “staked on neoliberal

economic principles” (van Dijck, 2013). In economic terms, the emergence of new digital technologies, which include social media platforms, has enabled value to be captured and created in new ways (Kenney & Zysman, 2015). Platforms ultimately “monetize human effort” by collecting massive amounts of data from their users based on their profiles, communications, content, and behavior (Fuchs, 2015; Kenney & Zysman, 2016). Users are the commodity, and their data is what generates profit for platform companies through advertising revenue (Fuchs, 2015). Whether or not platform companies intentionally frame themselves as democratic tools, they are undoubtedly sites of global power and influence. As Kenney and Zysman note:

We are in the midst of a reorganization of our economy in which the platform owners are seemingly developing power that may be even more formidable than was that of the factory owners in the early industrial revolution. (Kenney & Zysman, 2016)

However, platform companies do not yield power in terms of their economic capital alone. Platforms are not only run by corporate actors; they are sites of social exchange. The economic interests of platform companies therefore also influence the digital architecture and user guidelines of their products and services.

Content policies and user guidelines of platforms are another dimension of platforms influenced by capitalist logic. As Gillespie notes, these guidelines are also an example of why platform neutrality is a myth. Platform companies are always making decisions about what to leave up and what to take down, and often—although not always—these decisions align with companies’ economic interests. In-platform advertisements, for example, highlight how companies ensure “privileged real estate” and visibility to their financial partners (Gillespie,

2018). Highlighting or prioritizing advertisements is therefore both an architectural and financial decision made by platform companies that aligns with capitalist logic.

The act of *enforcing* these content policies and community guidelines also points to the ways in which market interests influence platform companies' content moderation practices, or lack thereof; the very fact that platform companies are companies therefore affects what content that appears on their products and services. Platform companies like Facebook, Twitter and YouTube serve billions of users across the globe and are tasked with moderating content on a massive scale. To date, the largest platform companies have historically relied on human moderators and artificial intelligence, or some combination of the two. Human moderators typically review content flagged by users, while artificial intelligence is deployed to automatically remove certain types of media, such as "terrorist propaganda" and hate speech. However, Gillespie calls these moderation efforts "imperfect solutions" (Gillespie, 2018). There will always be content that slips through the cracks, and you cannot retroactively remove whatever impact that content might have had. The biggest platform companies now serve billions of users across the globe and do not employ nearly enough human reviewers to effectively remove all problematic content. Furthermore, artificial intelligence cannot be deployed in contexts where there isn't a large enough dataset for an algorithm to accurately identify and remove content (Perrigo, 2019). Automated systems may also mistakenly remove content that doesn't violate a platform's user guidelines, again highlighting the imperfect nature of moderation, at scale.

Content moderation is also an expensive endeavor, and major platform companies outsource their content moderation efforts to minimize overhead costs by hiring contractors.

While the average Facebook employee makes \$240,000 annually, some of the company's contracted moderators made only \$28,000 a year, according to a 2019 investigation (Newton, 2019). Facebook has also been criticized for not hiring enough moderators who speak local languages. When the company received negative press coverage for not adequately responding to and moderating hate speech and anti-Rohingya propaganda in Myanmar, civil society pointed to a lack of Burmese-speaking content moderators, a failing that Mark Zuckerberg himself acknowledged (Thomas, 2019). From a critical perspective, Facebook's shortcomings can be attributed to a desire to maximize profit and minimize costs. Facebook's association with the Rohingya genocide also indicates that platforms are linked to human rights discourse and human rights-related content in more ways than one.

Legal and Policy Landscape

Discussions of platform accountability are also linked to the legal and policy landscape in which platforms are situated. While platform companies must comply with a number of laws, Section 230 of the Communications Decency Act is perhaps the most relevant within the context of this research project. Written in 1996, the Section 230 states that "no provider or user of an interactive computer service shall be treated as a publisher or speaker of any information provided by another information content provider." This means that, with few exceptions, internet service providers, including platform companies, cannot be liable for the user-generated content that they host (Kosseff, 2019). Section 230 was therefore instrumental in creating the social and legal framework in which social media platforms currently sit. Content moderation at scale is nearly impossible given the massive amounts of user-generated content uploaded to sites like Twitter, YouTube, and Facebook. Section 230 shields platform companies from potential

liability, while largely enabling these companies to implement their own user guidelines and community standards (*Section 230 of the Communications Decency Act*, n.d.). These guidelines and standards effectively become private forms of governance on speech and expression, influencing the ways in which we all interact and communicate.

Governance by Platform Companies

It is clear that platforms have been instrumental in disseminating information globally. Furthermore, it is evident that the economic and legal landscape in which platform companies operate shape the architecture and content policies of social media platforms. In recent years, scholars have also argued that this mix of content policy and platform architecture constitutes a form of privatized governance. While a more traditional articulation of governance refers to state power and “the government’s ability to make and enforce rules, and deliver services,” international relations scholars have expanded the word’s meaning to connote “how global life is organized, structured, and regulated” (Fukuyama, 2013; Solomon, 2006). Therefore, conceptualizing platform companies’ massive power as a form of global governance acknowledges the ways in which they organize, structure, and regulate speech and information online (Gorwa, 2019b).

While there is a large body of scholarly work on internet governance, researchers have only recently begun to explore the ways in which platform companies specifically constitute their own unique form of global governance. Denardis and Hackl, for example, survey the ways in which platform companies are governing anonymous speech, individual privacy, expression, and interoperability of their users by analyzing these platforms’ technological affordances, user policies, and corporate transparency data. They conclude that, due to the widespread global use

of their products and services, social media platforms have become “central points of control” online. From this perspective, platforms have effectively become conduits for and arbiters of online expression (DeNardis & Hackl, 2015). Platform companies therefore yield more than economic power alone; they have the power to shape whether and how information flows on a global scale.

Platform Accountability and Content Governance

A growing acknowledgement of the cultural and economic power that social media companies yield has also been met with a call for these companies to engage in more socially responsible business. Platform accountability refers to efforts to hold platform companies accountable amidst concerns regarding the misuse of their products and services (Philip Verveer, 2019). These concerns are numerous. Hate speech, incitement of violence, algorithmic bias, data breaches, terrorist propaganda, and misinformation are just some of the issues, controversies and scandals that have enveloped platform companies in recent years. This research paper is concerned with platform accountability within the particular context of content governance.

Content governance refers to decisions made about content moderation and enforcement, including whether and how content is removed or amplified on social media platforms (Pallero & Pírková, 2020). The United Nations, civil society, and a number of legal scholars have all argued that platform companies should take a human rights-centric approach to content governance, for multiple reasons. First, companies have a responsibility to respect human rights as outlined in the UN Guiding Principles on Business and Human Rights (*Guiding Principles for Business and Human Rights*, 2011). In addition, as “private stewards of public space,” UN Special Rapporteur David Kaye argues that international human rights law could also lend a sense of legitimacy to

content moderation practices (Kaye, 2019) Finally, as Access Now, a technology and human rights advocacy organization argues, content moderation practices have a “disparate impact” on those most at risk of human rights violations (Pallero & Pírková, 2020). All of these arguments, taken together, outline the case for a human rights-based approach to content governance, and this paper goes onto examine how civil society actors in particular make this case.

Platform Accountability Efforts via Multistakeholderism

While the issue of platform accountability poses its own unique challenges, the question of how to hold a multinational corporation accountable in the absence of traditional or legally binding governance mechanisms is not a new one. In other industries, non-binding, voluntary governance commitments or initiatives have emerged as one way to set standards that better hold companies accountable. These commitments and initiatives, which bring together actors from across different sectors of society, are referred to by Raymond and DeNardis as a form of multistakeholderism.

Multistakeholderism refers to “two or more classes of actors engaged in a common governance enterprise concerning issues they regard as public in nature, and characterized by polyarchic authority relations constituted by procedural rules” (Raymond & DeNardis, 2015). In part, because the term has been used so broadly, they argue that multistakeholderism tends to function as more of a “buzzword rather than an identifiable institutional form.” Still, this reality doesn’t delegitimize multistakeholderism; instead, it simply reveals that there are multiple variations of it. Moreover, Raymond and DeNardis outline four specific actors who partake in multistakeholderism: states, formal intergovernmental organizations (IGOs), firms, and civil society actors (Raymond & DeNardis, 2015).

Today, there are multiple identifiable forms of multistakeholderism within the context of platform accountability. They include Global Internet Forum to Counter Terrorism (GIFCT), The Christchurch Call, Santa Clara Principles on Content Moderation (SCP), The Twitter Trust and Safety Council (TWC), and Facebook's Oversight Board, to name a few (Gorwa, 2019b). While these initiatives vary in their makeup and orientation, most of them bring a number of civil society actors to the table to engage in some form of standard setting or governance process with platform companies and other sectors of society.

Civil Society, Corporate Accountability, and Social Change

Civil society organizations can also be part of platform accountability efforts outside the context of multistakeholderism. In business ethics and management literature, stakeholder engagement refers to *any* corporate practice or activity involving stakeholders (Dawkins, 2014). The formalized rules and processes of multistakeholderism are not a prerequisite.

Furthermore, civil society can contribute to platform accountability efforts without engaging directly with platform companies at all. With regards to corporate accountability challenges in other industries, civil society organizations have mobilized around issues that include child labor, deforestation and pollution. They may rely on confrontation, criticism and media coverage to bring issues to light, a strategy typically referred to as “naming and shaming” (Utting, 2005).

However, civil society is not uniform. Within the context of transnational social learning, or the “processes that increase awareness, capacities, and repertoires of action,” civil society organizations may serve a variety of different roles (Brown & Timmer, 2006). Nonprofit organization scholars Brown and Timmer identified five roles that civil society organizations can

serve in social learning: identifying issues, facilitating voice of marginalized stakeholders, amplifying the importance of issues, building bridges among diverse stakeholders, and monitoring and assessing solutions (Brown & Timmer, 2006). If platform accountability is a process that involves social learning, then civil society actors concerned with content governance and human rights may be expected to serve some if not all of these roles.

Platform accountability can also be conceptualized as a form of social change. Battilana and Kimsey, who study organizational behavior and social change at Harvard Business School, argue that those who participate in movements for social change can play three different roles: agitator, innovator and orchestrator. As they note:

An agitator brings the grievances of specific individuals or groups to the forefront of public awareness. An innovator creates an actionable solution to address these grievances. And an orchestrator coordinates action across groups, organizations, and sectors to scale the proposed solution. Any pathway to social change requires all three. Agitation without innovation means complaints without ways forward, and innovation without orchestration means ideas without impact. (Battilana & Kimsey, 2017)

Battilana and Kimsey offer another useful framework for beginning to understand how civil society organizations—and for the purposes of this paper, human rights advocacy groups in particular—are involved in processes of content governance and platform accountability.

Civil Society and Platform Accountability

With particular regards to platform governance and accountability, international relations scholar Robert Gorwa notes that civil society actors have served an important watchdog function through their advocacy and research efforts (Gorwa, 2019b). Civil society organizations are

ultimately part of a larger governance triangle that also involves state and corporate actors, and they have been pivotal in setting content governance standards such as the Santa Clara Principles and the Manila Principles on Intermediary Liability. As Gorwa notes, both of these standards and processes were established with little to no involvement from state or corporate actors (Gorwa, 2019b).

Civil society organizations ultimately serve as an important counterweight to both corporate and state power. Regardless of whether platform companies are noble in their attempts to serve the public interest, they are ultimately obligated to generate profits and lack both independence and representativeness (Gorwa, 2019b). Within the context of platform accountability, and in comparison to other informal governance processes, there are also more state interests at play regarding standards setting for online content; government pressure to crack down on “terrorist” or violent extremist content serves as one example (Gorwa, 2019b). Furthermore, we cannot always trust governments to act in their citizens’ best interests or to protect human rights (MacKinnon, 2012). Therefore, if social media platforms are to be sites of participation, free expression and democratic exchange, then civil society organizations can serve as a proxy for the global public or publics that platform companies serve. Given the global reach and impact that platform companies yield, in addition to the human rights value that content posted to social media platforms offer, human rights organizations in particular are especially important in ensuring that international human rights norms, values and interests are not overlooked or compromised in light of state or corporate interests. Closer examination of these organizations is therefore warranted.

Gorwa points to the need to establish a shared research agenda and methodology in order to evaluate the legitimacy of platform companies and the various informal governance initiatives that involve them. He points to three areas of inquiry: *legitimation politics*, *actor competencies*, and *power relations* (Gorwa, 2019b). This research project examines how human rights organizations understand their competencies and experience power relations in the realm of platform accountability and content governance.

Section III: Methodology

Section III provides an overview of this paper's theoretical frames, research questions, research methods, and processes of data collection. This section also briefly touches on the scope of this research project and the inherent limitations.

Theoretical Frames

The review of scholarly literature reveals that understanding the different dimensions of platform accountability requires both an inter- and multidisciplinary approach. For the purposes of this paper, conceptualizing the power that social media platform companies yield—economically, politically and socioculturally, or otherwise—is rooted in critical and constructivist theories of mass media. Platform companies are media institutions that act as central points of control, but individuals can also use social media platforms in different contexts and circumstances. Moreover, this paper assumes that content governance and platform accountability efforts are *sites of struggle* where large platform companies, governments, organized civil society, and the individual users that platforms serve compete for ideological hegemony (*A Dictionary of Media and Communication*, n.d.)

This project also draws on the disciplines of political science and international relations in addition to the work of legal scholars and prominent digital rights figures to understand how platform companies enact their own forms of governance. Finally, this paper relies on frameworks established in organizational and social movement literature to understand how various actors can contribute to social change. In particular, research data will be evaluated with reference to Battilana and Kimsey's *agitate-innovate-orchestrate* framework and the five roles Brown and Timmer identify that civil society actors play within the context of transnational social learning (Battilana & Kimsey, 2017; Brown & Timmer, 2006).

Research Questions

Three research questions were developed to explore how human rights organizations understand their actor competencies and experience power relations with regards to platform accountability and content governance.

- **Question 1:** What *strategies* do civil society organizations use in order to advocate for human rights within the context of content governance and platform accountability?
- **Question 2:** How do these organizations *understand* their role in ongoing human rights advocacy and platform accountability efforts, relative to platform companies?
- **Question 3:** What major *challenges* and *opportunities* do human rights experts and civil society organizations identify when it comes to holding platform companies accountable to their human rights impacts?

Given this project's theoretical framework, questions were designed to allow for a qualitative inquiry into how civil society actors experience power relations and understand their competencies within the platform accountability and content governance space.

Research Methods

Due to the qualitative nature of this research project, I chose to rely on semi-structured, in-depth interviews as my primary method of data collection. Interviews allow for greater inquiry into context, meaning, experiences, emotions and values. Given the nature of my research questions and their concern with both power dynamics and actor competencies, interviews were a particularly attractive research method.

Data Collection

Data was collected through a series of seven semi-structured interviews conducted in March 2020. Four of these interviews were with members of organized civil society, all of whom work for international human rights organizations and have been part of some stakeholder engagement process with social media platform companies in the past. Three of these interviews were with business and human rights professionals: experts who work or who have worked for consulting firms focused on human rights and sustainability. All three of these experts have worked directly with a platform company or have facilitated some process of stakeholder engagement in the past.

Informants in this study were recruited in one of two ways: through existing relationships or through referral. How each informant was recruited is disclosed in the table below. Potential informants were sent a brief overview of my research project in advance, which included a few short contextual paragraphs and my initial research questions. Informants were also notified in advance that all interviews would be anonymized and that no reference would be made to their particular organizations.

Interviews took place either in person or remotely by video conference and were audio recorded with informant consent. Informants were emailed a list of preliminary questions in advance, although most interviews became more open and unstructured discussions, as intended. To preserve the anonymity of each participant, gender neutral pronouns were also used. Table I provides an overview of each informant and interview.

Table I

Informant	Sector	Date of Interview	Method of Recruitment
Informant 1	Civil society	03 March 2020	Existing relationship
Informant 2	Civil society	18 March 2020	Referral
Informant 3	Civil society	17 March 2020	Referral
Informant 4	Civil society	27 March 2020	Referral
Informant 5	Consultant	05 March 2020	Existing relationship
Informant 6	Consultant	24 March 2020	Existing relationship
Informant 7	Consultant	25 March 2020	Existing relationship

Once data collection was complete, the audio recording of each interview was transcribed. I utilized inductive coding to identify key themes from each interview. Transcripts were coded three times. First, I coded each transcript by question type. Second, I coded each transcript by theme and unit of analysis. Finally, I cross-referenced these first two transcripts and

assigned each theme to a particular question in order to organize my findings and analysis sections accordingly.

Scope and Limitations

This research project and methodology has its inherent limitations. The short span of this research project limited the number of interviews I could conduct. Given the range of positions that civil society actors can take, it is expected that a number of perspectives will be overlooked in my analysis. Furthermore, while human rights principles are universal, inalienable, indivisible, interdependent and interrelated, different civil society actors within the human rights space may each focus their advocacy efforts on different rights, norms, and principles. The four members of organized civil society I spoke with also all know each other and have worked together in the past. Furthermore, these organizations are all situated closely to one another in terms of the human rights issues they work on and advocate for. Platform accountability and content governance brings a wide range of civil society actors to the table, their interests and advocacy efforts varied, and the diversity of informants interviewed could have been improved. However, this research project offers a starting point for evaluating how human rights organizations understand actor competencies and power relations in the platform accountability and content governance space. Their experiences and insight may echo those of other civil society actors and can be situated within existing literature.

Section IV: Findings

Section IV discusses the findings and key themes that emerged from the data as they pertain to the three initial research questions.

Question 1

What strategies do civil society organizations use in order to advocate for human rights within the context of content governance and platform accountability?

Question 1 is concerned with the *strategies* for human rights advocacy and the *nature* of engagement between civil society and social media platform companies. The data suggests, perhaps unsurprisingly, that there is no uniform stakeholder engagement *process*; the nature of engagement between civil society organizations and platform companies can vary; interacting with platform companies is largely context-specific. However, two key strategies did emerge: applying *external pressure* and leveraging *private networks*.

All four civil society actors remarked that their organizations used both of these strategies. Each strategy is distinct and comes with its own opportunities and constraints; however, both strategies leverage a company's inherent aversion to risk and incentive to maximize profits. This is, of course, a critical observation, but it is one that both consultants and members of civil society made. Furthermore, data suggests that this critical observation informs these organizations' *practical* approaches to human rights advocacy with regard to platform accountability and content governance.

External Pressure

The civil society actors I spoke with sought to increase platform accountability by applying external pressure through negative press and media coverage. All four informants from civil society identified this traditional "naming and shaming" strategy as the most effective and consistent lever in getting platform companies to meaningfully address their human rights impacts, or at the very least, listen to and engage with their organizations. Consultants also spoke

to the impact of external pressure on increasing platform accountability, although they did so with less frequency. This strategy is ultimately a form of agitation that relies on platform companies being reactive to reputational and financial risk.

Throughout the interviews, civil society spoke at great length about the role external pressure plays in platform accountability and stakeholder engagement. Their remarks were made both in the abstract and with particular regards to their own experiences with platform companies. One informant, a researcher at a human rights organization concerned with content takedowns, is quoted saying that “press has really helped in putting pressure on platforms... I think press is the biggest lever we have” (Informant 2). Similarly, another informant remarked that increased accountability within platform companies, in general, has been the result of “civil society working its ass off to draw attention to what these companies are doing, hand-in-hand with the press... without press attention, it’s very difficult to get these companies to move” (Informant 4). Informants from civil society repeatedly identified the press not only as a key lever, but as the main lever, throughout their interviews.

In some instances, informants provided specific examples of how the press has played a key role in getting platform companies to engage with their particular organizations. One informant credits the press coverage on content takedowns and the “PR disaster” that this created as a key motivation for a platform company agreeing to speak with their organization. This informant was even able to name the particular journalist and publication that they believe had the biggest impact (Informant 2). Two other informants believe that they are interacting more consistently with platform companies because they are now “regularly quoted in the press” and have authored op-eds on content enforcement in well-known media publications (Informant 3,

Informant 4). It is clear that the civil society actors interviewed explicitly identify the press as a necessary lever for increasing platform accountability and stakeholder engagement. The specificity and frequency with which these informants spoke of press coverage also affirms their confidence in and reliance on this particular strategy of human rights advocacy with platform companies.

Opportunities and Constraints of External Pressure

Both an opportunity and constraint, informants identified that, as a human rights advocacy strategy, applying external pressure on platform companies works because it takes advantage of a business' aversion to risk and incentive to maximize profits. While this is a critical observation, it is one that every informant made and every member of civil society in particular recognized as a method for getting platform companies to either address their human rights impacts or engage with external stakeholders on some level. Every informant explicitly mentioned the need to consider the role that shareholders play. As one informant, a researcher at a well-known human rights advocacy organization, noted:

[Civil society] needs to remember that these companies are all publicly listed companies, and the goal of a publicly traded company is to maximize shareholder profit, so if maximizing shareholder profit doesn't mean defending human rights and doing the right thing ethically and having a human rights based approach to your platform, then you're not going to care... I think to get these organizations to change you have to start hurting them in the pocket (Informant 3)

Similarly, another informant described the press as a major lever that civil society can apply, at scale, because it has the potential to “devastate the tech company in terms of their reputation”

which ultimately poses a “real threat to their bottom line” (Informant 1). It is evident that civil society understands the critical forces that motivate platform companies, and as a practical advocacy strategy, seeks to capitalize on these forces by applying external pressure.

However, civil society remarked that reliance on the press to generate external pressure can also *constrain* their advocacy efforts. As one informant notes, “without press attention, it's very difficult to get companies to move” (Informant 4). Interestingly, civil society also identified Facebook’s *particular* dominance in the news cycle as a key constraint that prevents the same amount of media coverage to be afforded to other companies. The same informant goes on to remark:

It's also really easy to get attention from the press on Facebook. It's harder to get the press to focus on YouTube. When you just don't have press attention on something, it's hard to get the companies to change. (Informant 4)

Similarly, another informant remarks that “Facebook has a bad reputation, but other platforms have gotten away with not having such a bad reputation” (Informant 3). Both of these informants were quick to caveat that, in many ways, this press coverage is warranted; Facebook’s accountability challenges and human rights impact deserve coverage. However, their frustrations suggest that Facebook’s media dominance, in effect, limits media coverage on other platform companies, hindering similar advocacy and accountability efforts. As another informant notes, “when we talk about social media, we're largely talking about Facebook” (Informant 2). These comments not only point to practical frustrations but also signal that, both rhetorically and symbolically, Facebook has more or less become the surrogate for all of the problems associated

with social media platforms; this ultimately limits the press attention that other platform companies, also a constraint of the external pressure model of advocacy.

Private Networks

In addition to applying external pressure through media coverage, informants revealed that they also engage *directly* with social media companies for human rights advocacy purposes through various forms of stakeholder engagement. While the nature of these engagements and relationships was much more varied, along with the degree of specificity with which they were discussed, each informant from civil society either had working relationships with employees within platform companies or had been part of some sort of stakeholder engagement process in the past. Generally speaking, however, all four of the organizations I spoke with relied, at least in part, on private networking, relationship-building, and behind-the-scenes stakeholder engagement processes for human rights advocacy purposes. The variety and multiplicity of these relationships is an observation that business and human rights consultants also identified.

The first informant, an international criminal justice expert, spoke of their experience working closely with another civil society organization to facilitate a “closed-door event” to discuss human rights-related content with platform companies. This informant also identified RightsCon, an annual summit on human rights in the digital age, as a “really important hub” for building relationships and contacts within companies (Informant 1). A second informant revealed that its organization had developed useful contacts within platform companies, although they noted that these contacts have changed over time (Informant 2). Another informant revealed that they had attended a meeting with Facebook and a number of other civil society organizations in the past, but noted that it has historically been “different people reaching out” from within the

same platform company (Informant 3). A fourth informant notes that human rights advocacy with platform companies requires “painstaking relationship building” (Informant 4). It is evident that civil society organizations not only rely on external pressure to advocate for human rights but also leverage private networks, although the nature of these networks clearly vary.

Informant *perceptions* of this particular strategy also varied across civil society organizations. One informant remarked that their relationship with platform companies “ebb and flow,” although they also called their relationship with a particular platform company “nice and proactive” (Informant 4). Another informant characterized their organization’s relationship with platform companies as much more reactive. This informant commented that platform companies “contact us when they need something or when they want to talk, but any product changes just happen without consultation” (Informant 3). These comments, of course, are general statements, and the value of this particular inquiry does not rely on whether or not civil society organizations categorize their engagement with platform companies as “good” or “bad.” However, the range of perceptions civil society expressed regarding the nature of their relationships does suggest that the private networks model of advocacy is much more complex and context-dependent than the external pressure model.

The data, taken in the aggregate, reveals no uniform approach to or perception of stakeholder engagement processes. It is possible that this is a result of the study’s small sample size and the lack of specificity with which these relationships were discussed, however, the multiplicity of these relationships and stakeholder engagement processes is an observation that business and human rights consultants also made. In reference to the technology industry, at large, one consultant remarked that “there is no cohesive external engagement strategy... there

needs to be better coordination on what external engagement looks like for these companies” (Informant 6). Another business and human rights professional, who has facilitated stakeholder engagement processes with civil society and technology companies in the past, commented that, today, there are “so many different places where companies interact with [civil society], so many places for dialogue. It’s a rich, rich field” (Informant 7). Their comments suggest that the variety of private networks reflected in the data may very well be representative of an industry-wide trend, or lack thereof, with particular regards to the nature of stakeholder engagement.

Opportunities and Constraints Within Private Networks

Although the data suggests that the nature and form of private networks are, again, both varied and context-specific, every informant from civil society still acknowledged that the reality of business was a key constraint that affects how human rights advocacy takes place within private networks. However, civil society also acknowledged the presence of so-called “good actors” inside companies—those who care about human rights, who are sympathetic to civil society’s concerns, and who understand the impetus for a human-rights based approach to business. Within the context of human rights advocacy via private networks, three out of four informants identified the need to build and leverage relationships with internal actors who are willing to listen and who can facilitate change from within institutions. One informant, whose organization is concerned with content takedowns, pointed to the role these internal actors have played in helping them get content reinstated:

Behind the companies are people, and in general, the people that I've spoken to, all of them from these platforms, they get the issue. They understand why it's important. They would like to find some solution. But they're limited because the problem is so big and

the structures in which they work in [have] too many factors at play. But that doesn't mean they don't do anything. That doesn't mean they don't care... you know, it's not like a good versus bad... there's a lot of gray areas... in some cases, they've been extremely helpful going like above and beyond what they've needed to do. You know, as people, because they really get it. (Informant 2)

Here, the informant acknowledges both an opportunity and constraint. They express gratitude for the help of internal actors who are sympathetic to civil society's grievances and who, in some cases, are doing more than they normatively "should" be by assisting civil society organizations to get content reinstated. However, the informant also acknowledges that these internal actors are still limited and must operate within existing corporate structures to orchestrate or facilitate change.

Similarly, a different informant touched on how identifying these internal actors can be another approach to human rights advocacy and an avenue for communicating civil society's concerns to platform companies:

Every one of these corporations has people who don't care about the human rights consequences and care about profit above all, and [these corporations also] have people who do care. And so, you know, you have to find those people and talk to them... if there are people at the company who want it to be less dangerous for people and who want it to be a better place for free expression and [who] care about human rights, then I'm happy to be the person who, sort of, you know, bites the bullet and talks to them. (Informant 4)

Again, this informant identifies the crucial role internal actors can play in advocating for human rights within companies, and within corporate structures that, from a critical perspective, exist solely to generate profit.

Data suggests that human rights advocacy, by way of private networks, requires a much more nuanced approach than applying external pressure. Both a constraint and opportunity, another key theme that emerged from the data was that leveraging private networks requires an understanding of how business operates. Informants discussed this reality either within the context of identifying internal actors who could drive change, as previously described, or, with regards to the need to make an explicit “business case” for human rights. This latter point was also reaffirmed by business and human rights consultants.

One informant from civil society recalls their experience learning to speak “corporate” at the time they began to engage with platform companies on a more regular basis:

I didn't realize what a learning curve I was going to have to go through to be able to speak corporate. Because I don't come from a business background, or business and human rights, and I didn't understand the frames through which [businesses] look at the world and the ways that they think and the way they speak. (Informant 1)

This informant's comments suggest that human rights advocacy through direct engagement with platform companies is constrained by business objectives and culture. This is another insight with both critical and practical dimensions. It reveals underlying power dynamics, skewed in the favor of platform companies, but it also constitutes a strategy for human rights advocacy via private networks.

Question 2

How do these organizations understand their role in ongoing human rights advocacy and platform accountability efforts, relative to platform companies?

Question two is more theoretical and reflexive than the first research question. It is an attempt to explore how civil society organizations view their *normative* role in platform accountability efforts and stakeholder engagement processes and is an inquiry into how these organizations understand human rights advocacy within the context of platform companies. While many of the findings relate back to the results of the first research question, question two enables greater analysis of the institutional power dynamics at play.

Several findings emerged. First, informants acknowledged that civil society is not uniform but that each organization can encourage platform accountability in its own way. Second, although the civil society organizations interviewed acknowledged that they rely on both external pressure and private networking as avenues for human rights advocacy, none of these organizations expressed any concerns in regards to balancing these two strategies. Civil society organizations did, however, acknowledge the privilege they yield in having access to direct engagement with large platform companies and express frustrations over the lack of stakeholder engagement with smaller, local organizations and human rights activists.

Civil Society's Normative Role

All four informants from civil society suggested that within the context of platform accountability and human rights advocacy, civil society organizations ultimately serve as a check on corporate power. One informant asserted that it is their organization's role, normatively

speaking, to “keep rattling the cage” (Informant 3). The motivations for doing so have been largely discussed with regards to question one.

However, all four informants also acknowledged, either explicitly or implicitly, that each civil society organization serves its own purpose and takes its own approach to advocacy. Every business and human rights consultant made similar claims. As one informant put it, “each civil society group has a unique position in the space and they look at the world through their lens” (Informant 1). This is a finding that informants also implicitly affirmed by acknowledging the work that other civil society organizations have done with regards to platform accountability. One informant, for example, commented that the Electronic Frontier Foundation (EFF) has been a “big pioneer in tech company advocacy,” noting that the organization’s transparency reports have been largely influential in improving accountability within the technology industry (Informant 4). Throughout the interviews, every other civil society informant also praised or at least mentioned the work that other organizations were doing with regards to platform accountability. Furthermore, while the organizations interviewed all relied on both external pressure and private networks as methods of human rights advocacy, they also acknowledged that approaches to advocacy exist on a spectrum. This is a point that each business and human rights professional reaffirmed.

Tensions and Power Dynamics Across Roles

Findings under Question 1 reveal that organized civil society advocates for human rights in at least one of two ways—by applying external pressure or by leveraging private networks. Findings from the first question also establish that, with regards to *both* strategies, the business

interests of platform companies must be considered. Unsurprisingly, organized civil society was also acutely aware of the immense power and influence these companies yield and understand that it is part of their normative role to apply some form of external pressure to platform companies. However, informant responses suggest that applying external pressure, for the most part, has not affected their ability to also leverage private networks or participate in stakeholder engagement processes as part of their advocacy strategy.

Informants did acknowledge that there is inherent tension to applying external pressure while also leveraging private networks, in part due to power dynamics skewed in the favor of platform companies. However, none of these informants viewed this tension as a major concern. One informant notes:

It's a tricky fine line between rattling the cage and being welcomed to the table, and obviously, you don't want to get yourself disinvited from the table, but equally, we have to be able to speak truth to these organizations... So is this a challenge? No, everyone knows your role as [redacted]... They might not like what we said, they might have gotten pissed off by what we said, but at the same time, there's never been an issue... We've said what we've said and there's never been an issue with that. (Informant 3)

This informant's response indicates that although their organization may criticize platform companies, this criticism has not practically affected any working relationship. Similarly, another informant notes that balancing criticism and collaboration is not an issue, because their organization is ultimately creating value for platform companies:

[We don't just] come to companies to say you're doing this wrong. [Redacted] also comes before there's a problem... We are doing them a service. We're also doing a service for

human rights defenders. But it's basically like free consultation... It's really been in their business interests to listen. And there's not much friction in that relationship, even when we're critical... I've said very harsh things about companies publicly, and I will continue to do so. On the other hand, you know, I come with solutions. (Informant 4)

Informant responses suggest that, at least to some extent and in some cases, it is within the interests of platform companies to not only listen to external criticism but to also maintain relationships with these civil society groups. From the perspective of civil society, these relationships seem to be worth maintaining so long as organizations can continue to engage in public dialogue and public criticism.

Power Dynamics Between Civil Society

Informants from organized civil society, all of whom were part of large and well-known human rights organizations, also acknowledged the power dynamics present *between* civil society organizations and how this impacts the ability of smaller, local organizations to engage directly with platform companies. As one informant noted, platform companies tend to listen to and engage with civil society organizations based on their reputation and impact: “They listen to organizations based on reputation, and that have a global reputation, so it's always easier for global NGOs to talk to them.” All civil society actors interviewed echoed this sentiment. This particular informant goes on to note that reputation is important, in part, because platform companies have a hard time vetting smaller, local organizations (Informant 4). However, the effect is that large, well-funded global NGOs dominate stakeholder engagement processes. Another informant recalls attending a meeting at Facebook to discuss the company’s human rights impact in Myanmar, despite not being incredibly knowledgeable on the region. They

comment, “Am I really the right person to be speaking about Myanmar? Some ways yes, some ways no. I didn’t say much in that meeting because I felt like a complete imposter and fraud” (Informant 3). These comments suggest that global civil society organizations not only have greater opportunity to engage directly with platform companies but that actors within these organizations are aware of the privilege they yield.

Given their elite position, informants from civil society also considered it part of their role to either facilitate connections between local organizations and platform companies or amplify the voices, grievances, and concerns of local organizations and human rights activists. As one informant notes:

I’m in a very privileged position. And I have to think about what that means. And I have to think, what are the moments every single moment where I can say, ‘Actually, you should be talking to this person.’ One of the major roles that I see myself having in this field broadly actually is both encouraging my colleagues that are a better resource for organizations and to help companies talk to the right people. (Informant 4).

Another informant also reflects on the privilege that they yield, being part of a well-known organization based in California and having direct contacts some of the major platform companies:

There’s a bit of a Western Northern elite privilege operating. And I think it’s incumbent on those of us who do have access to use that access in ways that are sensitive to understanding that there are many groups who don’t. And can we bring some of those perspectives to the front? (Informant 1).

Similarly, a third informant remarks that they believe part of their role is to serve as a “middleman” between local activists and journalists and the large platform companies whose decisions impact these individuals (Informant 2). Informant responses reveal that power dynamics between civil society organizations in large part determine whether and how these organizations engage directly with platform companies or have access to private networks.

Question 3

What challenges and opportunities do human rights experts and advocates identify in the ability for social media companies to meaningfully address their human rights impacts?

Question three discusses the challenges and opportunities related to platform accountability and human rights advocacy that civil society actors and business and human rights experts explicitly identified throughout their interviews. The biggest challenge that civil society identified was a platform company’s *follow-through* on their human rights obligations and commitments. Given the business interests of platform companies, which has been thoroughly discussed in this paper’s literature review and with regards to the first two questions, this insight is in many ways unsurprising. However, a key opportunity that all informants identified was the emergence of human rights teams and human rights leads *within* platform companies, many of whom come from civil society. While many of these informants expressed cautious optimism, they acknowledged that these teams can act as internal advocates and have the potential to facilitate the follow-through that was previously identified as a challenge.

Challenges

The biggest challenge that organizations identified with regards to human rights advocacy and platform accountability was turning any sort of stakeholder engagement process into meaningful action. One informant referred to a meeting at a particular platform company as nothing more than a “brain pick” in which the company “cherry picked” which recommendations to implement and which to ignore (Informant 3). Similarly, another informant notes, “Well, I think there's a difference between listening and acting... We've seen a lot of willingness to listen, but we haven't seen so many changes being made.” (Informant 2). What these comments ultimately reveal is a larger key challenge that civil society organizations face within the context of platform accountability at large: corporations are not obligated to listen to them or implement their recommendations. This represents a major limitation in the competency of civil society actors within informal governance spheres and is indicative of a larger power imbalance. Absent state regulation, platform companies can ultimately decide whether they wish to implement more human rights-sensitive policies and practices. If these practices don't clearly align with business interests, or are not articulated in a way that align with business interests, then the incentive for a platform company to implement recommendations made by civil society is low.

Opportunities

However, every civil society actor and business and human rights professional also identified a key opportunity for human rights within the platform governance space. Platform companies are beginning to hire human rights teams and human rights leads, and these internal actors present an opportunity for greater stakeholder collaboration with human rights advocacy groups. In addition, these internal actors represent the presence of a human rights perspective

that is institutionally embedded *within* companies and may be able to facilitate the change and follow-through that is currently missing. As one informant notes:

There's been a huge sort of hiring process of people from civil society by Facebook and Twitter... I have actually had a friend, like someone I would call a friend, who works at Facebook, which I could not have said two years ago. You're really seeing this crossover from NGOs into companies (Informant 4).

Similarly, another informant acknowledges that this particular moment in time, where platform companies are beginning to hire human rights teams, is largely unprecedented:

I feel like we're at a moment in time. It's really fascinating. And I would say it started about six months ago and it's kind of picking up speed... all of a sudden a lot of people I work with on the humanitarian side and the human rights side are all going corporate and they're becoming insiders. (Informant 1).

These hiring decisions, the movement of human rights professionals from civil society into platform companies, emerged as a major point of discussion in every interview, both with civil society organizations and consultants. They represent another dimension of platform accountability.

With regards to their *perception* of these new internal human rights teams, and whether or not they would be effective, civil society exercised cautious optimism. They argued that these teams have the *potential* to advance human rights within companies, however, this requires both adequate funding and resources. As one informant noted, regardless of whether or not these hiring decisions were made in good faith or as a PR stunt, what matters is the ability of these employees to ultimately drive change. As they note:

I think some of those people really care.. and if the companies are only doing it because they get bad press like whatever, that's the reality. That's capitalism. You know how it works. So whether they're doing it because it's the right thing to do or they're doing it because they've gotten so much pressure, they are doing it. It's happening... I think it is really important when companies hire people to try to understand if they're actually getting resources... I think that's actually sometimes part of the strategy for people in civil society, as well, is to encourage those people to have more resources. (Informant 4)

Similarly, another informant remarked that when evaluating the potential for these internal actors to drive change, evaluating where their funding is coming from is also important (Informant 3). If these teams are not funded, or if they do not have buy-in from company leadership, then ultimately, these teams will wind up toothless. This is an observation and caveat that all members of civil society and all business consultants made.

Still, the general sentiment that civil society actors communicated was one of cautious optimism. Knowing some of these hires personally, they viewed their intentions as legitimate and the potential for change as strong, of course, given adequate funding and buy-in. Interviews ended on largely optimistic notes, and when asked what has driven platform accountability efforts to date, one informant remarked:

I think I think it's really the result of pressure. It's the result of the press. And now, increasingly, I would say it's going to be the result of some people who are actually, you know, trying to make the change from the inside (Informant 4).

Another informant, discussed how this “new moment” presents a key opportunity for platforms to step up their accountability efforts:

I think they're starting to get it, slowly. That doesn't mean they're going to get it. That doesn't mean they're going to do the right thing, but I think there is an opportunity where they might do the right thing (Informant 3).

What these comments ultimately suggest is that constraints on human rights advocacy have shifted and new opportunities have emerged. Human rights employees within platform companies have the potential to serve as internal human rights advocates, which can help address some of the key challenges identified by civil society. First, these employees can translate human rights language into “corporate,” a key constraint identified in regards to the first question. Second, these employees can help create more uniform stakeholder engagement strategies, addressing another constraint identified in Question 1.

Furthermore, these employees have the potential to increase the diversity of stakeholders that companies reach out to and bring in more local organizations, another issue that civil society organizations identified as discussed with regards to Question 2. Finally, these actors can turn listening into action, overcoming a key challenge that civil society explicitly has explicitly identified. Of course, these outcomes are dependent on the organizational culture of each company and the resources these internal teams receive, but the potential impact, given findings from the first two research questions, is evident.

Section V: Discussion

Section V situates the findings of this study within existing literature. It discusses the relevance of these findings and identifies areas for further research. Moreover, these findings,

when situated within frameworks of social learning and social change, illuminate potential gaps that civil society could fill to potentially improve platform accountability efforts.

Relevance

Three identifiable roles emerged from the data that align with transnational social learning. They are: *identifying issues*, *amplifying the importance of issues*, and *facilitating voice of marginalized stakeholders*. In addition, using Battilana and Kimsey's framework, the organized civil society actors in this study primarily identified their role as *agitators*. While the role of *orchestrating* social change may soon rest on internal human rights leads, there is still greater opportunity for civil society and companies to act as *innovators*.

Study Findings and Social Learning

As discussed in this paper's findings, civil society organizations can advocate for human rights by applying external pressure via media coverage and leveraging private networks. These strategies align with two of the five roles that civil society actors can play in transnational social learning processes, as conceptualized by Brown and Timmer: *identifying issues* and *amplifying the importance of issues*. Findings suggest that civil society both identifies and amplifies the importance of issues via media coverage, although more proactive private relationships between platform companies and civil society organizations may enable these organizations to directly identify and present issues to companies without relying on media coverage.

This study's findings also echoed a third role identified by Brown and Timmer: *facilitating the voice of marginalized stakeholders*. As discussed with regards to Question 2, the civil society organizations interviewed, all more well-resourced human rights organizations, recognized their privilege and how this privilege enabled a relationship between them and

platform companies. These organizations also considered it part of their normative role to act as “middlemen” or to point platform companies in the direction of local organizations. In other words, these civil society organizations were *facilitating the voice of marginalized stakeholders*. Mapping this study’s findings onto Brown and Timmer’s framework suggests that civil society’s account is consistent with existing literature.

Study Findings and Social Change

Findings were also consistent with Battilana and Kimsey’s theory of social change, which outlines three roles that are present in movements for social change: *agitators*, *innovators*, and *orchestrators*. Aligning research findings with this model suggests that civil society primarily serves as an agitator. The organizations interviewed understood that it was their normative role to “rattle the cage” and hold platforms to account, primarily by applying external pressure through media coverage.

Findings suggest that civil society also has the potential to play the role of *innovator*, however, this role is limited, in part by the business interests of platforms. Being part of the process to create actionable solutions in large part requires that civil society have a relationship with the platform companies they wish to change or better align with human rights values. While civil society does engage in advocacy by leveraging private networks, as outlined in Question 1, the challenges discussed in Question 3 suggest that platform companies have been slower to act than civil society would like.

Data does not suggest that civil society organizations have served as orchestrators within the realm of platform governance. Orchestrators coordinate action across groups, organizations, and sectors (Battilana & Kimsey, 2017). However, this is a role that human rights leads have the

potential to fill given their unique position. Situating this study's findings within Battilana and Kimsey's framework can therefore offer an explanation as to why certain actionable change has not taken place as discussed in Question 3. If social change is to be realized, then all three roles need to be filled.

Further Research

Based on this study's findings, two areas of further research stand out. Given that this research project relied on insights from larger civil society organizations, and given that these organizations themselves pointed to the lack of stakeholder engagement with smaller, local organizations, further research into how these local organizations understand content governance and platform accountability is warranted. Furthermore, civil society actors pointed to the emergence of these new internal human rights teams as potential opportunities for increased human rights accountability and norm setting within platform companies. However, as new phenomena, these roles and their effectiveness have yet to be studied. Studying how corporate human rights actors, within platform companies, is therefore a second site of potential inquiry.

Conclusion

This research project attempted to better understand the competencies of civil society actors and the power dynamics at play within platform accountability efforts and did so by focusing on three areas of inquiry: strategies, normative roles, and challenges and opportunities. This study was also conducted through theoretical frames of critical theory of mass media and social movement literature. Findings are consistent with prior literature in both of these disciplines. While stakeholder engagement and strategies of human rights advocacy varied, informants both explicitly and implicitly identified the role that platform companies play as

corporate actors and how their profit motives may ultimately limit or complicate human rights advocacy efforts. However, civil society also strategically deployed these corporate interests within the context of private networking, signaling that this project has both critical and practical value. Civil society actors understand the profit motives of platform companies, and in some cases, can use this to create value that may sometimes align with human rights-related interests. Still, it is evident that power dynamics were at play and also influenced human rights advocacy efforts in more ways than one.

Situating this project's findings within frameworks of social change and social learning not only validate that the roles civil society play within content governance and platform accountability efforts are consistent with other research but also offers insight as to where platform accountability can be improved, adding another practical dimension to this project. For example, if meaningful social change is to be realized within the platform accountability space, then civil society cannot rely on agitation alone. Innovators and orchestrators must also be part of the equation. Perhaps the emergence of internal human rights teams will fill this role, however, critical analysis highlights why this approach can be limited.

The critical dimensions of this study are perhaps more important than ever. They reveal that corporate interests are always part of the equation, even as platform companies build out their own governance processes and despite the growing number of human rights actors within companies. While independent bodies such as Facebook's Oversight Board may signal the rise of a new "Process" era of content governance, many critics question whether these processes can ever be truly independent. Platform companies may act as governments, but they ultimately still lack the democratic accountability of public bodies. Until this governance gap is addressed,

organized civil society must serve as a proxy for the global public or publics that platform companies serve, advocating for decisions and processes that respect human rights and democratic norms. Increasingly, civil society organizations must also serve as a check on the very processes their advocacy efforts have helped establish. Lines may blur, processes may change, and approaches may evolve, but holding platform companies accountable still requires some sort of independent check on corporate power, a role that organized civil society must continue to fill.

Works Cited

A Dictionary of Media and Communication. (n.d.). Oxford Reference.

<https://doi.org/10.1093/oi/authority.20110803100509105>

Abdul Rahman Al Jaloud, Hadi Al Khatib, Jeff Deutch, Dia Kayyali, & Jillian C. York. (2019).

Caught in the Net: The Impact of “Extremist” Speech Regulations on Human Rights Content. Electronic Frontier Foundation.

Allen, M. (2013). What was Web 2.0? Versions as the dominant mode of internet history. *New*

Media & Society, 15(2), 260–275. <https://doi.org/10.1177/1461444812451567>

Battilana, J., & Kimsey, M. (2017). Should You Agitate, Innovate, or Orchestrate? *Stanford*

Social Innovation Review.

https://ssir.org/articles/entry/should_you_agitate_innovate_or_orchestrate

Bogost, I., & Montfort, N. (2009). Platform Studies: Frequently Questioned Answers. *University*

of California, Irvine, 7.

Bowers, J., & Zittrain, J. (2020). Answering impossible questions: Content governance in an age

of disinformation. *Harvard Kennedy School Misinformation Review*, 1(1).

<https://doi.org/10.37016/mr-2020-005>

Brown, L. D., & Timmer, V. (2006). Civil Society Actors as Catalysts for Transnational Social

Learning. *Voluntas: International Journal of Voluntary and Nonprofit Organizations*, 17(1), 1–16. <https://doi.org/10.1007/s11266-005-9002-0>

Carter, C. A. (2016). Police Brutality, the Law & Today’s Social Justice Movement: How the

Lack of Police Accountability Has Fueled #Hashtag Activism. *CUNY Law Review*, 20(2), 521–558.

- Dawkins, C. E. (2014). The Principle of Good Faith: Toward Substantive Stakeholder Engagement. *Journal of Business Ethics*, 121(2), 283–295.
<https://doi.org/10.1007/s10551-013-1697-z>
- DeNardis, L., & Hackl, A. M. (2015). Internet governance by social media platforms. *Telecommunications Policy*, 39(9), 761–770. <https://doi.org/10.1016/j.telpol.2015.04.003>
- Dubberley, S., Koenig, A., & Murray, D. (2020). *Digital Witness: Using Open Source Information for Human Rights Investigation, Documentation, and Accountability*. Oxford University Press.
- Establishing Structure and Governance for an Independent Oversight Board. (2019, September 17). *About Facebook*. <https://about.fb.com/news/2019/09/oversight-board-structure/>
- Fuchs, C. (2015). *Culture and Economy in the Age of Social Media*. Routledge.
- Fukuyama, F. (2013). What Is Governance? *Governance*, 26(3), 347–368.
<https://doi.org/10.1111/gove.12035>
- Gillespie, T. (2010). The politics of ‘platforms.’ *New Media & Society*, 12(3), 347–364.
<https://doi.org/10.1177/1461444809342738>
- Gillespie, T. (2017). Governance of and by platforms. *SAGE Handbook of Social Media*.
<http://culturedigitally.org/2016/06/governance-of-and-by-platforms/>
- Gillespie, T. (2018). *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media*. Yale University Pres.
<https://yalebooks.yale.edu/book/9780300173130/custodians-internet>
- Gorwa, R. (2019a). What is Platform Governance? *Information, Communication & Society*, 22(6), 854–871. <https://doi.org/10.1080/1369118X.2019.1573914>

- Gorwa, R. (2019b). *The platform governance triangle: Conceptualising the informal regulation of online content*. Internet Policy Review. <https://doi.org/10.31235/osf.io/tgnrj>
- Guiding Principles for Business and Human Rights*. (2011). United Nations Office of the High Commissioner for Human Rights. <https://www.unglobalcompact.org/library/2>
- Helmond, A. (2015). The Platformization of the Web: Making Web Data Platform Ready. *Social Media + Society*, 1(2), 2056305115603080. <https://doi.org/10.1177/2056305115603080>
- Hoffmann, A. L., Proferes, N., & Zimmer, M. (2018). “Making the world more open and connected”: Mark Zuckerberg and the discursive construction of Facebook and its users. *New Media & Society*, 20(1), 199–218. <https://doi.org/10.1177/1461444816660784>
- Kaye, D. (2019). *Speech Police: The Global Struggle to Govern the Internet*. Columbia Global Reports.
- Kenney, M., & Zysman, J. (2015, June). *Choosing a Future in the Platform Economy: The Implications and Consequences of Digital Platforms*. Kauffman Foundation New Entrepreneurial Growth Conference, Amelia Island, Florida.
- <https://web.archive.org/web/20180704124615/http://www.brie.berkeley.edu/wp-content/uploads/2015/02/PlatformEconomy2DistributeJune21.pdf>
- Kenney, M., & Zysman, J. (2016). The Rise of the Platform Economy | Issues in Science and Technology. *Issues in Science and Technology*, 32(3).
- Kosseff, J. (2019). *The Twenty-Six Words That Created the Internet*. Cornell University Press.
- MacKinnon, R. (2012). *Consent of the Networked: The Worldwide Struggle For Internet Freedom*. Basic Books, Inc.
- Newton, C. (2019, February 25). *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>

Pallero, J., & Pírková, E. (2020). *26 recommendations on content governance: A guide for lawmakers, regulators, and company policy makers*. Access Now.

<https://www.accessnow.org/guide-how-to-protect-human-rights-in-content-governance/>

Perrigo, B. (2019, November 26). Facebook Says It's Removing More Hate Speech Than Ever Before. But There's a Catch. *Time*.

<https://time.com/5739688/facebook-hate-speech-languages/>

Philip Verveer. (2019). *Platform Accountability: An Interim Measure*. Harvard Kennedy School Shorenstein Center on Media, Politics, and Public Policy.

<https://shorensteincenter.org/platform-accountability-interim-measure/>

Raymond, M., & DeNardis, L. (2015). Multistakeholderism: Anatomy of an inchoate global institution. *International Theory*, 7(3), 572–616.

<https://doi.org/10.1017/S1752971915000081>

Romm, T. (2019, October 17). Zuckerberg: Standing For Voice and Free Expression.

Washington Post.

<https://www.washingtonpost.com/technology/2019/10/17/zuckerberg-standing-voice-free-expression/>

Section 230 of the Communications Decency Act. (n.d.). Electronic Frontier Foundation.

Retrieved May 17, 2020, from <https://www.eff.org/issues/cda230>

Shirky, C. (2011). The Political Power of Social Media: Technology, the Public Sphere, and

- Political Change. *Foreign Affairs*, 90(1), 28–41. JSTOR.
- Solomon, M. S. (2006). Power and the Politics of Global Governance. *International Studies Review*, 8(2), 327–329. JSTOR.
- Srnicek, N. (2016). *Platform Capitalism*. Polity Press.
- <https://www.wiley.com/en-us/Platform+Capitalism-p-9781509504862>
- Thomas, E. (2019, June 21). Facebook Keeps Failing in Myanmar. *Foreign Policy*.
- <https://foreignpolicy.com/2019/06/21/facebook-keeps-failing-in-myanmar-zuckerberg-ar-akan-army-rakhine/>
- Tufekci, Z., & Wilson, C. (2012). Social Media and the Decision to Participate in Political Protest: Observations From Tahrir. *Journal of Communication*, 62(2), 363–379.
- Utting, P. (2005). Corporate responsibility and the movement of business. *Development in Practice*, 15(3–4), 375–388. <https://doi.org/10.1080/09614520500075797>
- van Dijck, J. (2013). *The Culture of Connectivity: A Critical History of Social Media*. Oxford University Press.
- <https://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199970773.001.0001/acprof-9780199970773>